

Université de Montréal

**Mécanisme de régulation des gènes par le récepteur des
oestrogènes alpha : fonction des éléments de réponse des
oestrogènes (ERE) en tant qu'enhancers distaux.**

par

Julie Deschênes M.Sc.

Département de Biochimie, Institut de Recherche en Immunologie et Cancérologie
Faculté de Médecine

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de doctorat
en Biochimie



Août 2007



© Julie Deschênes, 2007

W
4
U58
2008
v.015

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

Mécanisme de régulation des gènes par le récepteur des oestrogènes alpha : fonction des
éléments de réponse des oestrogènes (ERE) en tant qu'enhancers distaux.

présentée par :

Julie Deschênes M.Sc.

a été évaluée par un jury composé des personnes suivantes :

Martine Raymond PhD, président-rapporteur

Sylvie Mader PhD, directeur de recherche

Alain Verreault PhD, membre du jury

Josée Dostie PhD, examinateur externe

Résumé

Ayant de nombreuses cibles tissulaires, les oestrogènes jouent des rôles diversifiés chez l'humain. Leur principal rôle au niveau cellulaire est la régulation de la prolifération. Les récepteurs des oestrogènes (ER) appartiennent à la grande famille des récepteurs nucléaires stéroïdiens, et agissent par liaison directe à l'ADN. Ils reconnaissent un élément de réponse palindromique de 15 pb (ERE). Les données cristallographiques indiquent que la liaison à l'ADN ne se fait pas exactement comme celles des autres récepteurs nucléaires. L'étude exhaustive des acides aminés responsables de la liaison aux deux nucléotides centraux de l'ERE a démontré qu'il est possible de changer la spécificité de liaison du récepteur en remplaçant ces acides aminés. Bien que cette spécificité et la sélectivité du ER envers l'ADN soient connues et que plusieurs gènes cibles des oestrogènes soient identifiés, peu de ERE de haute affinité sont répertoriés. La recherche de ces ERE dans le génome de la souris et de l'humain a établi qu'ils sont distribués dans l'ensemble du génome, avec un enrichissement dans la région proximale de l'initiation de la transcription des gènes. Le nombre élevé de ces ERE soulève la question de la fonctionnalité *in vivo* de ces éléments. Des gènes cibles primaires des oestrogènes ont été identifiés dans une lignée de cellules dérivée d'un carcinome mammaire en présence d'un inhibiteur de la synthèse protéique. Un fort enrichissement des ERE prédits se retrouve à proximité des gènes primaires positivement régulés par les oestrogènes. L'étude du promoteur du gène GREB1, fortement exprimé dans les tumeurs mammaires et semblant jouer un rôle dans la prolifération cellulaire, a démontré que les trois ERE consensus situés dans le promoteur peuvent collaborer entre eux pour médier la transcription du gène, malgré une grande distance (20 K pb) les séparant. Des études de boucles de chromatine ont permis de développer un modèle de coopération entre les ERE afin de participer à la transcription de ce gène en présence d'oestrogènes. L'ensemble de ces études permet de comprendre à la fois les mécanismes de liaison du ER sur l'ADN et démontre que la régulation des gènes par les oestrogènes peut se faire via des éléments lointains.

Mots-clés : œstrogène, récepteur nucléaire, ER α , ERE, gène, transcription, MCF7, GREB1, boucle de chromatine.

Abstract

Estrogens play pleiotropic roles in their different target tissues in the human body. At the cellular level, one of their major effects is to induce cellular proliferation. The estrogen receptors (ERs) are members of the steroid nuclear receptor family. Once bound to estrogens, they effect their action through direct binding to a specific 15bp palindromic DNA sequences (EREs). The crystallographic data ER-ERE complexes demonstrated intricate interactions in the central base pairs of the ERE. Our study of two amino acids important for this binding demonstrated that it is possible to alter the ER DNA binding in the specificity changing the identity of these amino acids. However, although the selectivity of the receptors for DNA is now well characterized and several target genes are identified, only a few EREs have been characterized previously. Our screening for high-affinity EREs in the entire mouse and human genome demonstrated a distribution of EREs throughout the genome, with an enrichment in the vicinity of the transcriptional start sites of genes. The high-frequency of EREs in the genome raised the question of their functional role *in vivo*. Estrogen primary target genes in breast carcinoma cell lines were identified by expression microarrays in presence of a protein synthesis inhibitor. An enrichment in EREs was found in the vicinity of estrogen primary up-regulated genes. Finally, we studied the promoter of the primary target gene GREB1, a gene expressed in breast tumors and involved in cell proliferation. We demonstrated that its three consensus EREs, spread over 20 Kbp of promoter region, cooperate for the transcriptional regulation of GREB1 by estrogen. Chromatin loop studies suggested a multi-partite model of enhancer for its transcription. All these studies increased the understanding of the mechanisms involved in ER DNA binding and demonstrated long-range regulation of transcription by ERs.

Keywords : estrogen, nuclear receptor, ER α , ERE, gene, transcription, MCF7, GREB1, chromatin loop.

Table des matières

RÉSUMÉ	III
ABSTRACT	IV
TABLE DES MATIÈRES	V
LISTE DES TABLEAUX	VIII
LISTE DES FIGURES	IX
LISTE DES SIGLES ET ABRÉVIATIONS	XII
REMERCIEMENTS	XV
PREMIÈRE PARTIE	1
CHAPITRE 1 : LES OESTROGÈNES : HORMONES STÉROÏDIENNES	2
1.1 SYNTHÈSE ET DÉGRADATION	2
1.2 RÔLES ET FONCTIONS PHYSIOLOGIQUES	5
1.2.1 Le tractus reproducteur	5
1.2.1.1 Chez l'homme	5
1.2.1.2 Chez la femme	7
1.2.2 Les glandes mammaires	9
1.2.3 Les tissus osseux	10
1.2.4 Le système nerveux	11
1.2.5 Le système cardiovasculaire	12
1.3 LES MALADIES ASSOCIÉES AUX OESTROGÈNES	12
CHAPITRE 2 : RÉCEPTEUR DES OESTROGÈNES, MÉDIATEUR DE LA SIGNALISATION CELLULAIRE	15
2.1 LES RÉCEPTEURS NUCLÉAIRES DES OESTROGÈNES	16
2.1.1 Dualité de fonction : récepteur et facteur de transcription	16
2.1.2 L'activation et les modifications post-transcriptionnelles	19

	vi
2.1.3 La liaison à l'ADN	21
2.1.4 La distribution tissulaire	23
2.1.5 Les récepteurs orphelins reliés aux oestrogènes (ERR)	24
2.2 LA SIGNALISATION DES ŒSTROGÈNES	25
2.2.1 Signalisation nucléaire classique	26
2.2.1.1 Les cofacteurs de la signalisation nucléaire	26
2.2.1.2 Recrutement ordonné des complexes impliqués dans la signalisation classique	27
2.2.2 Signalisation nucléaire alternative	29
2.2.3 Signalisation membranaire : mode d'action rapide des oestrogènes	30
CHAPITRE 3 : CHROMATINE ET TRANSCRIPTION DES GÈNES	33
3.1 L'ORGANISATION SPATIO-TEMPORELLE DE LA CHROMATINE	33
3.1.1 Les composantes des promoteurs	34
3.1.2 Structure tridimensionnelle et dynamique de la chromatine	36
3.2 RÉGULATION DE LA TRANSCRIPTION À LONGUE DISTANCE	40
3.2.1 La mise à jour du concept de promoteur	41
3.2.2 Les boucles de chromatines	43
CHAPITRE 4 : PROBLÉMATIQUES	47
DEUXIÈME PARTIE	49
CHAPITRE 5 : GENOME-WIDE IDENTIFICATION OF HIGH AFFINITY ESTROGEN RESPONSE ELEMENTS IN HUMAN AND MOUSE	50
CHAPITRE 6 : MECHANISMS OF GENE REGULATION BY ESTROGEN RECEPTOR ALPHA IN BREAST CANCER CELLS.	106
CHAPITRE 7 : REGULATION OF GREB1 TRANSCRIPTION BY ESTROGEN RECEPTOR ALPHA THROUGH A MULTIPARTITE ENHANCER SPREAD OVER 20 KB OF UPSTREAM FLANKING SEQUENCES	157

TROISIÈME PARTIE **180**

CHAPITRE 8 : DISCUSSION ET PERSPECTIVES	181
8.1 FIXATION À L'ADN DE ER À L'ÉCHELLE DU GÉNOME	182
8.1.1 La prédiction de site de fixation de ER <i>in silico</i>	183
8.1.2 Les sites de liaison prédits sont-ils liés <i>in vivo</i> ?	185
8.2 GÈNES DIRECTEMENT RÉGULÉS PAR ERA	186
8.2.1 ER α médiateur de la signalisation oestrogénique pour tous les gènes régulés?	187
8.2.2 Les rôles fonctionnels des gènes régulés directement par ER	188
8.3 MÉCANISTIQUE DE LA TRANSCRIPTION SUR DE LONGUES DISTANCES	190
8.3.1 Les ERE distaux sont impliqués dans la formation de boucle de chromatine.	191
8.3.2 Fonctionnalité des boucles de chromatine dans les glandes mammaires	194
8.4 L'INTÉGRATION DES RÉSEAUX DE RÉGULATION DE LA TRANSCRIPTION	195
8.4.1 Coopérativité de ER α avec d'autres facteurs de transcription	195
8.4.2 L'amplification du signal oestrogénique	197
8.4.3 Etablissement de boucles de rétroaction	198
8.5 CONCLUSION	199

BIBLIOGRAPHIE **201**

**ANNEXE 1 : RATIONAL DESIGN OF AN ESTROGEN RECEPTOR MUTANT WITH
ALTERED DNA BINDING SPECIFICITY** **XVI**

Liste des tableaux

Deuxième Partie

Chapitre 5

Table 1 :	Estrogen response elements in the human genome and corresponding elements in the mouse genome.	p. 88
Table 2 :	EREs with identical sequences in human and mouse orthologs.	p. 89
Table 3 :	EREs found in upregulated estrogen target genes.	p. 90
Table s1 :	EREs found in both human and mouse orthologs.	p. 96

Chapitre 6

Table 1 :	Top hundred genes regulated by E2 in the presence or absence of cycloheximide.	p.151
Table s1 :	Transcription factor binding sites enriched in a 5 or 10 kb window around the transcriptional start sites of significantly regulated genes.	p.154
Table s2 :	Transcription factor binding sites enriched in a 5 or 10 kb window around the transcriptional start sites of the top 100 regulated genes in each category.	p.155

Liste des figures

Première Partie

Chapitre 1

- Figure 1 : La synthèse des oestrogènes dans le follicule ovarien. **p. 4**

Chapitre 2

- Figure 2 : Les récepteurs des oestrogènes. **p. 18**
Figure 3 : Interaction de ER α avec l'ADN. **p. 23**
Figure 4 : Modèle de signalisation dynamique et directe de ER α . **p. 28**

Chapitre 3

- Figure 5 : Le promoteur minimal eucaryote. **p. 34**
Figure 6 : La structure de la chromatine. **p. 37**
Figure 7 : La technique 3C : *Chromosome Conformation Capture*. **p. 43**

Deuxième Partie

Chapitre 5

- Figure 1 : *In vitro* binding of ERs to consensus and near-consensus EREs. **p. 92**
Figure 2 : Distribution of EREs in the human and mouse genomes. **p. 93**
Figure 3 : *In vitro* binding of ER α to EREs identified in this study. **p. 94**
Figure 4 : *In vivo* binding of human ER α to EREs in ER-positive breast cancer cells. **p. 95**

Chapitre 6

- Figure 1 : Role of ER α in mediating the effect of estradiol in MCF7 breast cancer cells. **p.142**
Figure 2 : Cycloheximide-sensitive and resistant E2 target genes in

	MCF7 breast cancer cells.	p.143
Figure 3 :	Expression of several genes induced only in the presence of CHX at 24 hrs is regulated in its absence at earlier times.	p.144
Figure 4 :	Several transcription factors in the E2F family are primary target genes of estradiol in MCF7 breast cancer cells.	p.145
Figure 5 :	Regulation by estradiol of both up- and down-regulated primary target genes is dependent on ER α .	p.146
Figure 6 :	Estrogen response elements are enriched in the vicinity of primary UP-regulated genes.	p.147
Figure 7 :	Distribution of EREs in a 50 kb window around estrogen target genes.	p.148
Figure 8 :	Identification of functional EREs through bioinformatics screening of flanking regions of primary target genes.	p.149
Figure 9 :	E2F response elements are enriched in secondary up-regulated genes.	p.150
Figure S1 :	Validation of the estrogen regulation of selected genes at 24 hrs stimulation.	p.152
Figure S2 :	Estradiol-induced proliferation of MCF-7 breast cancer cells is blocked by cycloheximide.	p.153
 <i>Chapitre 7</i>		
Figure 1 :	GREB1 is transcriptionally induced by ER α .	p.173
Figure 2 :	Consensus EREs recruit ER α , SRC-3, acetylated histones and phosphorylated polymerase II in the presence of estradiol.	p.174
Figure 3 :	Induction of chromatin loops associating enhancers and the GREB1 transcriptional start site by estradiol treatment.	p.175
Figure S1 :	GREB1 mRNA is induced by estradiol in ER α -positive cancer cells.	p.176
Figure S2 :	The strong and sustained induction of GREB1 by estradiol is	

a primary transcriptional event. **p.177**

Figure S3 : The consensus GREB1 EREs function as a multi-partite enhancer. **p.178**

Annexe 1

Figure 1 : Model for the selectivity of steroid receptor/DNA interactions. **p. XVIII**

Figure 2 : Effect of replacements in both arms of the ERE palindrome on complex formation with full length ERs and isolated DBDs. **p.XXI**

Figure 3 : Chemical and steric constraints determine the specificity of complex formation with wt ER α . **p.XXII**

Figure 4 : ERs bind with high selectivity to PuGNNCA palindromes. **p.XXIII**

Figure 5 : Role of K210 and E203 in the affinity and specificity of receptor-DNA complex formation. **p.XXIV**

Figure 6 : Mutations of E203 do not generate expected switches in DNA binding specificity. **p.XXV**

Figure 7 : Mutation K210A introduced in E203 mutants reduces binding to the consensus ERE and reveals the switched specificity of mutant E203R. **p.XXVI**

Figure 8 : Altered patterns of reporter vector transactivation by ER mutants. **p.XXVI**

Figure 9 : Modeling of the effect of mutations at position 203 on interaction with palindromic response elements. **p.XXVII**

Liste des sigles et abréviations

α	:	alpha
AF	:	fonction activatrice 1 ou 2
AKT	:	protéine kinase B
ATP	:	adénosine tri-phosphates
β	:	bêta
BRE	:	<i>Transcription Factor II B Recognition Element</i>
cAMP	:	adénosine mono-phosphate cyclic
DBD	:	domaine de liaison à l'ADN
DES	:	di-éthystilbastro
DPE	:	<i>downstream promoter element</i>
E2	:	17 β -estradiol
EGF	:	facteur de croissance de l'épiderme
EGFR	:	récepteur du facteur de croissance de l'épiderme
eNOS	:	<i>Endothelial Nitric Oxide Synthase</i>
ER	:	Récepteur des oestrogènes (alpha et/ou bêta)
ERE	:	élément de réponse aux oestrogènes
ERR	:	Récepteur relié aux récepteurs des oestrogènes
FISH	:	hybridation <i>in situ</i> fluorescente
FSH	:	<i>folicule stimulating hormone</i>
GPR30	:	récepteur couplé aux protéines G-30
GR	:	récepteur des glucocorticoïdes
HAT	:	<i>histone acetyl transferase</i>
HDL	:	<i>high density lipoprotein</i>
IGF1	:	<i>Insulin-like Growth Factor 1</i>
IGF1R	:	<i>Insulin-like Growth Factor 1 Receptor</i>
IL	:	interleukine 1, 2, etc...
Inr	:	site d'initiation de la transcription

Ins	:	région frontière entre les gènes <i>silencer</i>
LBD	:	domaine de liaison du ligand
LCR	:	<i>locus control region</i>
LDL	:	<i>low density lipoprotein</i>
LH	:	<i>luteinizing hormone</i>
MAPK:		<i>mitogen activated protein kinase</i>
MCF7 :		<i>human breast adenocarcinoma cell line</i>
NF-kB:		<i>nuclear factor-kappa B</i>
Pb	:	paire de base (nucléotide)
PI3K	:	<i>phosphoinositide 3-kinases</i>
pS2/TFF1 :		<i>trefoil factor 1</i>
TAM	:	Tamoxifène
TBP	:	protéine liant la boîte TATA
TF	:	transcription factor TF II....
RNA-TRAP :		<i>tagging and recovery of associated proteins</i>
SFRE	:	<i>steroid factor response element</i>
Shc	:	<i>src homology 2 domain-containing</i>
SKBR3 :		<i>human breast cancer cell line</i>
SP-1	:	<i>transcription factor SP1 (specificity protein)</i>
SRC	:	<i>steroid receptor coactivator</i>
SWI/SNF :		famille de protéine reliée à SWI/SNF, <i>matrix-associated, actin-dependant regulator of chromatin</i>

À Simon et Charly...

Remerciements

Il est toujours difficile de faire des remerciements juste sans oublier personne. Je voudrais remercier ici les gens qui ont contribué de manière significative à ma réussite de quelques façons autant d'un point de vue scientifique que personnel. J'aimerais remercier les membres du jury pour avoir accepté de prendre un peu de temps pour faire de ma thèse une réussite.

En tout premier lieu, je voudrais remercier mes deux mentors qui ont fait de moi la scientifique que je suis. Premièrement, Dr André De Léan, mon superviseur de maîtrise qui m'a initié à la science. Je tiens à le remercier pour ses précieux conseils et son encouragement à la poursuite de mes études doctorales. Il m'a ouvert les yeux sur la science et m'a donné le goût d'élargir mes connaissances. Deuxièmement, Dre Sylvie Mader, ma directrice de thèse, qui a su m'aider à progresser à la fois dans la réflexion scientifique et la rigueur de la science. Elle m'a beaucoup aidé à développer ma persévérance, outil essentiel en science. Finalement, elle m'a guidé tout au long du développement technique et dans la poursuite de mon projet de thèse.

En deuxième lieu, une thèse n'est pas réalisable sans l'aide et le support de membres du laboratoire : Anick Auger M.Sc, Denis Nguyen M.Sc, Gisèle Marguier M.Sc, Élise Hébert M.Sc, David Ferland M.Sc et les autres membres du laboratoire. Chacun à sa façon a su m'aider à me dépasser, à atteindre mon but et à rendre agréable l'environnement de travail. Plus particulièrement, j'aimerais remercier la Dre Véronique Bourdeau, avec qui j'ai étroitement travaillé tout au long de ma thèse et qui m'a aidée jusqu'à la fin autant comme scientifique que comme guide. J'aimerais également remercier particulièrement Dr Walter Rocha, pour m'avoir à la fois aidée et conseillée d'un côté technique et m'avoir encouragé à persévérer.

En troisième lieu, j'aimerais souligner l'aide extérieur et les soutiens de ma famille, de mes amis(ies), particulièrement Dre Mélanie Lebel et le Jazz Daniel Big Band. Sans leurs supports et conseils ces dernières années auraient été difficiles.

Première partie

Chapitre 1 : Les oestrogènes : hormones stéroïdiennes

Chapitre 2 : Récepteur des oestrogènes, médiateur de la signalisation cellulaire

Chapitre 3 : Chromatine et transcription des gènes

Chapitre 4 : Problématiques

Chapitre 1 : Les oestrogènes : hormones stéroïdiennes

Le mot «hormone» provient du grec «*hormôn*» qui signifie exciter, mettre en marche ((1), <http://en.wikipedia.org>) ; les hormones transmettent un signal dans différents tissus chez les organismes multicellulaires, d'où le terme exciter. Bien que leur structure chimique soit diversifiée, le mécanisme d'action commun des hormones est de transmettre un message via un récepteur dans une cellule donnée. Elles sont généralement classifiées selon leur composition chimique (ex : stéroïdes, peptidiques) et leurs lieux primaires de synthèse. Les fonctions hormonales sont principalement : la reproduction et la différenciation sexuelle, le développement et la croissance, le maintien de l'homéostasie interne ainsi que la régulation du métabolisme et l'apport en nutriment (<http://www.ncbi.nlm.nih.gov/Literature/>).

Le mot «œstrogène» est dérivé du latin «*oestrus*» et -gène, signifiant «qui provoque l'oestrus ou cycle menstruel chez les femelles des mammifères» (1). Les oestrogènes appartiennent à la grande famille des hormones stéroïdiennes, qui sont dérivées du cholestérol sous l'effet de différentes enzymes spécifiques. Elles ont des effets physiologiques diversifiés ne se situant pas exclusivement dans le système reproducteur femelle. Ne différant pas en cela des autres hormones, les oestrogènes vont transmettre leur effets physiologiques via la liaison à différentes classes de récepteurs.

Avant d'entrer dans le détail des interactions des oestrogènes avec leurs différents récepteurs et des mécanismes moléculaires de ces interactions, voici un bref rappel sur la synthèse des oestrogènes et sur leurs principaux effets physiologiques.

1.1 Synthèse et dégradation

Les hormones stéroïdiennes sont des dérivés métaboliques du cholestérol. La conversion du cholestérol implique différents cytochromes P450 qui vont le transformer en androgènes, en progestines et en oestrogènes. Chez la femme, il existe deux principales

façons de synthétiser des oestrogènes : via une synthèse complète dans les ovaires (du cholestérol à l'estradiol en passant par l'aromatisation des androgènes) et via les androgènes circulants et/ou locaux (synthèse à partir de l'aromatisation des androgènes) (2-5).

Chez la femme, avant la ménopause, la majeure partie des oestrogènes est produite dans les ovaires. Ces hormones sont relarguées dans la circulation systémique pour médier leurs effets sur différents tissus. Cette production est contrôlée par les hormones hypophysaires LH et FSH. Ces hormones agissent sur l'expression et la régulation des P450 dans différentes cellules des ovaires. L'hormone LH agit via son récepteur de surface dans les cellules de la thèque situées dans le follicule. Il y a alors stimulation de la production d'oestrogènes, via l'augmentation des niveaux de la P450_{SCC} et de la P450_{17 α} (Fig. 1). La FSH, quant à elle, va plutôt agir sur les cellules granulaires du follicule pour augmenter la production de P450_{arom} (CYP19) et de 17 β -hydrohystéroïde déshydrogénase (Fig. 1) (4,5).

La production d'oestrogènes dans les cellules de la thèque est dépendante de la transformation du cholestérol en prégnenolone via la P450_{SCC} (Fig. 1). Par la suite, il y a transformations successives en progestérone et finalement en androsténone, le précurseur des différents oestrogènes. L'étape de transformation du cholestérol est limitante dans le processus de synthèse. La forme d'oestrogènes la plus active est l'estradiol. Du côté des cellules de la granulosa, il y a production d'oestrogènes via l'aromatisation des androgènes produits dans les cellules de la thèque (Fig. 1). Les métabolites sont ensuite relâchés dans le fluide folliculaire et peuvent ainsi rejoindre la circulation sanguine (4,5).

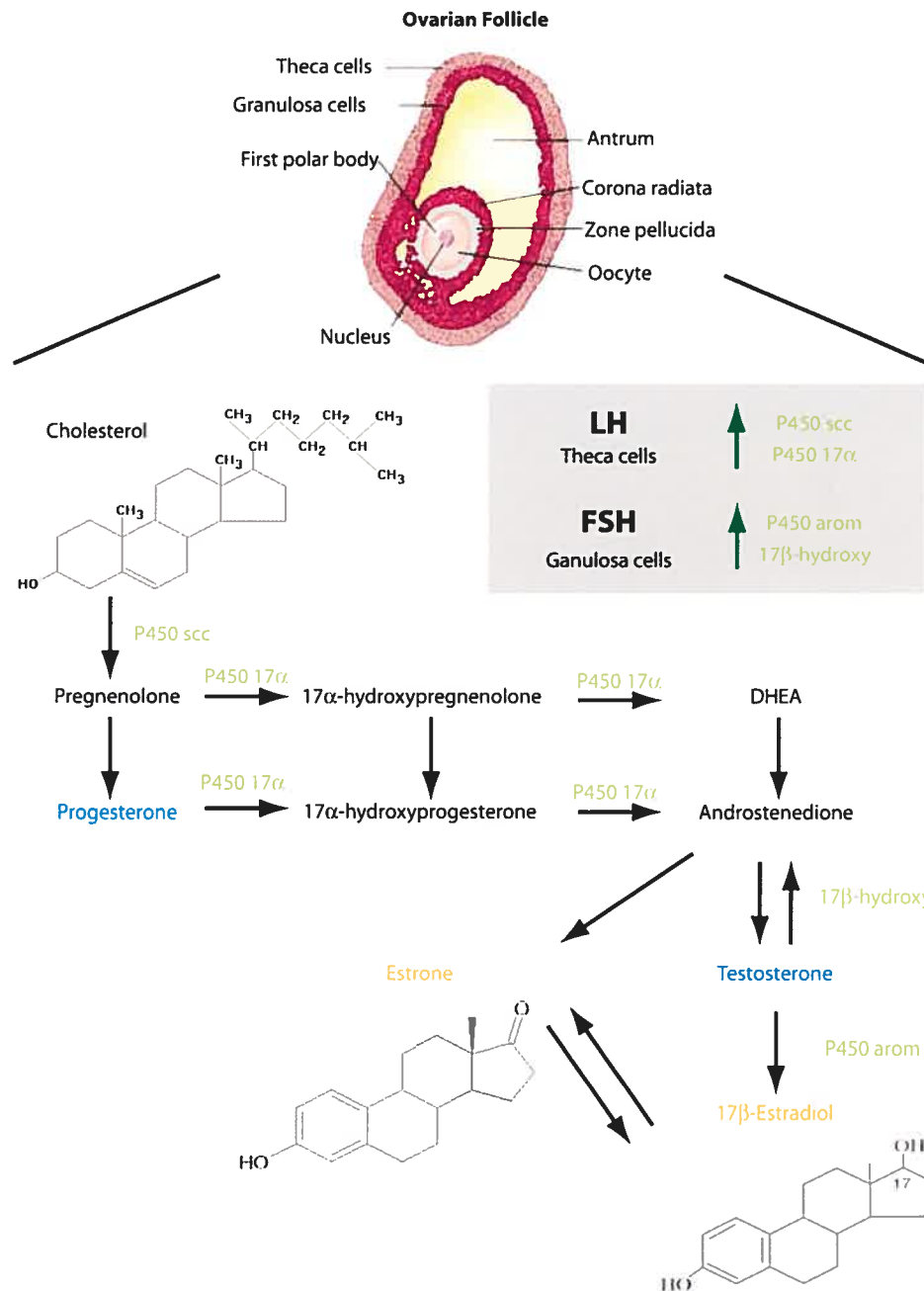


Figure 1 : La synthèse des oestrogènes dans le follicule ovarien. La synthèse des formes majeures des oestrogènes, le 17 β -oestradiol et l'oestrone, est représentée. Plusieurs combinaisons différentes de synthèses peuvent être effectuées pour produire les mêmes hormones. La progestérone ainsi que la testostérone peuvent entrer directement dans la voie métabolique. Adapté de Basic & Clinical Endocrinology (6).

Lors de la ménopause chez la femme, il y a arrêt de la production d'oestrogènes par les ovaires. Il reste donc seulement une production locale d'oestrogènes suivant l'aromatisation des androgènes circulants. Les tissus adipeux, les cellules mésenchymateuses, les ostéoblastes, l'endothélium des vaisseaux sanguins et différentes régions du cerveau ont recours à ce mode de production paracrine / intracrine. La production d'oestrogènes chez l'homme se fait aussi via l'aromatisation des précurseurs androgéniques particulièrement dans les testicules (4,5).

1.2 Rôles et fonctions physiologiques

Les oestrogènes ont été caractérisés à l'origine comme les hormones «femelles», jouant un rôle prépondérant chez la femme pour la reproduction. L'évolution de la science a permis de mieux comprendre les effets des oestrogènes dans les différents tissus chez la femme, mais également l'importance des oestrogènes chez l'homme, particulièrement dans le tractus reproducteur et les os. Les rôles majeurs des oestrogènes dans les différents tissus seront abordés dans la présente section.

1.2.1 Le tractus reproducteur

Les hormones stéroïdiennes jouent des rôles prépondérants dans les fonctions reproductrices. La logique voudrait que les oestrogènes soient importants chez la femme et que les androgènes le soient chez l'homme. La biologie est plus complexe, ainsi qu'il est apparu maintenant que le rôle des oestrogènes chez les deux sexes est de mieux en mieux connu.

1.2.1.1 Chez l'homme

Chez l'homme, les oestrogènes sont présents à faible concentration dans la circulation sanguine. Cependant, leur concentration dans le sperme et les fluides

testiculaires peut même dépasser ce qui est retrouvé dans la circulation chez la femme. La conversion locale des androgènes produits dans les testicules en oestrogènes explique ces concentrations, *a priori* surprenantes chez les mâles. Des doses pharmacologiques d'oestrogènes administrées à des mâles peuvent causer des problèmes au niveau des testicules et de la prostate. Chez les trans-sexuels, la thérapie par les oestrogènes à long terme va causer une hyperplasie des cellules de Leydig du testicule et une atrophie des canaux séminifères (7-9).

Une meilleure compréhension des effets locaux des oestrogènes a été possible avec les études sur des animaux *knock-out*. L'abolition des récepteurs des oestrogènes a permis de broser un tableau plus clair du rôle potentiel des oestrogènes chez les mâles. Les études physiologiques de ces animaux mâles ont démontré une diminution de la taille des testicules chez les mâles plus âgés. L'hyperplasie des cellules de Leydig a aussi été décrite comme concordant avec une augmentation des niveaux circulants de LH. Cette augmentation cause des niveaux deux fois plus élevés de testostérone (7,8,10). Il a aussi été démontré dans ces mêmes animaux mâles que les oestrogènes produits localement régulent la réabsorption des fluides luminaux dans la tête de l'épididyme (surface du testicule). En effet, il y a aussi apparition de dilatation de différentes régions du testicule. Ces effets s'aggravent fortement avec l'âge des animaux. Les défauts de réabsorption vont causer des effets menant à la réduction du compte de spermatozoïdes dans le sperme, et souvent à l'infertilité. Des études plus détaillées sur les causes de l'infertilité des animaux transgéniques ont également démontré des problèmes de maturation et de motilité des spermatozoïdes ainsi qu'une incapacité à fertiliser un oocyte *in vitro* (7,8,10).

Certaines études ont décrit des effets des oestrogènes dans les tissus accessoires du tractus reproducteur comme la prostate. Les oestrogènes semblent jouer un rôle dans la croissance, la différenciation et les activités fonctionnelles de l'épithélium de la prostate (9,11). Les études chez les animaux *knock-out* n'ont pas apporté de réponse claire à ce sujet (7,8,10). Il a cependant été démontré dans des cultures d'organes de rat que l'ajout d'une

forme d'oestrogène, le 17β -oestradiol (E2), augmentait la croissance de l'épithélium de ces organes (7).

Finalement, différents tissus du tractus reproducteur mâle expriment l'aromatase, une enzyme qui convertit les androgènes en oestrogènes. Cette enzyme a été détectée entre autres dans les cellules de Leydig, les cellules germinales et également dans le sperme humain (12). Il est maintenant reconnu que l'aromatisation se fait dans les cellules germinales de Leydig, de Sertoli, et l'épithélium de l'épididyme (11). Ceci vient renforcer la théorie d'un rôle des oestrogènes chez les mâles de différentes espèces.

1.2.1.2 Chez la femme

Bien qu'il existe plusieurs sites de synthèse des oestrogènes chez la femme, le site de synthèse principal est sans équivoque les ovaires. Les oestrogènes jouent différents rôles dans plusieurs tissus non-reproducteurs, cependant leur raison d'être principale réside dans le contrôle de la gestation et de la lactation. Il n'est donc pas surprenant que les oestrogènes aient un impact majeur sur les différentes composantes du système reproducteur des femmes (8).

Les tissus cibles des oestrogènes les mieux caractérisés dans le tractus reproducteur sont : les ovaires, les trompes de Fallope (ou oviductes), l'utérus, le col utérin et le vagin. Le développement des tissus reproducteurs chez les femelles se fait en deux grandes phases: une phase intra-utérine durant la vie embryonnaire (prénatale) et une phase de maturation lors de la puberté (postnatale). Les effets des oestrogènes semblent concentrés lors de la puberté et de la reproduction pendant la vie adulte. Ceci diffère des hormones stéroïdiennes masculines (androgènes) qui ont un impact majeur sur la différenciation durant la vie foetale chez l'homme. Il semble toutefois exister des différences entre les espèces puisque que des études d'infusion d'oestrogènes chez la souris et le rat ont démontré une augmentation du poids utérin seulement chez les individus femelles pré-pubères (8).

Les ovaires, lieu de synthèse des oestrogènes, sont aussi influencé par les niveaux d'œstrogènes locaux. Il existe principalement deux fonctions ovariennes: l'ovaire est un lieu de maturation du follicule qui va permettre, lors de l'ovulation, d'expulser un oocyte mature dans l'oviducte. C'est également le lieu de synthèse de différentes hormones stéroïdiennes et peptidiques. La synthèse d'œstrogènes est réalisée dans deux types de cellules : les cellules de la thèque (production à partir du cholestérol) et les cellules de la granulosa (production à partir des androgènes). Les œstrogènes se retrouvant dans les ovaires vont jouer un rôle important dans la maturation du follicule. Ils augmentent la prolifération des cellules de la granulosa et augmentent également l'action de la FSH et la synthèse du récepteur de LH qui est critique pour l'ovulation. Ultiment, les oestrogènes jouent aussi un rôle de régulation positive sur leur propre synthèse (8).

Les ovaires expulsent l'oocyte dans les oviductes, les canaux reliant les ovaires à l'utérus. Ils sont influencés par la production d'oestrogènes, menant à la prolifération des cellules de l'oviducte lors de l'augmentation de la concentration d'hormones régulée par le cycle menstruel. Il existe cependant des variations dans les effets de ces hormones suivant l'espèce et les doses d'oestrogènes utilisées (8).

L'utérus, relié aux oviductes, est également un lieu majeur d'action des oestrogènes. Servant à la nidification du blastocyste après la fécondation de l'ovule par un spermatozoïde, il sera le lieu de croissance du fœtus. Lors de la puberté chez la femme, les cellules utérines deviennent sensibles aux œstrogènes et commencent à connaître des cycles de prolifération et de différenciation suivant le cycle ovarien. Les oestrogènes contrôlent la prolifération cellulaire et mettent ainsi en marche la préparation de l'épithélium pour la réception possible d'un embryon à chaque cycle ovarien (ou menstruel). Cette préparation passe par l'activation de nombreux gènes et signaux servant à amplifier la réponse mitotique. Entre autres, l'augmentation de l'expression du récepteur de la progestérone est un facteur important lors de l'arrivée d'un embryon, dû au rôle de la progestérone lors de l'implantation (8).

Le vagin, qui est le conduit entre l'utérus et la partie externe de l'appareil génital, est également sensible aux oestrogènes. Durant le cycle menstruel, les cellules des muqueuses vaginales vont aussi subir une prolifération accrue et une différenciation dépendante des oestrogènes (8).

1.2.2 Les glandes mammaires

Lorsque l'on pense aux effets des oestrogènes chez la femme, la maturation des glandes mammaires et la différenciation lactogénique durant la grossesse sont les premiers effets qui viennent en tête. Les glandes mammaires possèdent plusieurs étapes de développement (13) et les oestrogènes interviennent de différentes façons lors de ces étapes. Lors de la vie fœtale, les oestrogènes ne semblent pas jouer un rôle prépondérant chez le fœtus féminin. Par contre, il est reconnu que les stéroïdes produits par les ovaires semblent jouer un rôle lors de la vie fœtale chez le rat (8).

Chez la femme, les effets critiques et marquants des oestrogènes arrivent lors de la puberté, de la grossesse et durant la vie adulte. Les oestrogènes sont impliqués dans la formation des bourgeons terminaux des conduits mammaires, la prolifération de l'épithélium des conduits et dans la formation des lobes alvéolaires durant la grossesse (8). Ils ont un rôle à jouer durant les nombreux cycles de prolifération du cycle menstruel à travers la vie adulte (14,15). Cependant, les mécanismes d'action précis des oestrogènes sont plutôt mal caractérisés (15,16). Il semble tout de même y avoir un consensus pour un double mode d'action de ces hormones : un premier aspect direct nécessitant une concentration faible d'oestrogènes et un deuxième aspect plus indirect, nécessitant une concentration hormonale plus élevée (15,17).

Le mode d'action direct des oestrogènes est présent à des concentrations basses d'hormone circulante et implique entre autres l'induction de l'expression du récepteur de la progestérone dans les cellules mammaires, l'augmentation du niveau de prolactine et la différenciation de l'épithélium (7,8,15). Il est à noter que la progestérone et la prolactine

jouent des rôles clé dans la formation des alvéoles et la production du lait par les glandes mammaires durant la grossesse (7,8,15).

Le mode d'action indirect (ou paracrine) passerait par l'augmentation de la concentration de l'EGF local dû à la production de ce facteur de croissance par les cellules stromales sous le contrôle des oestrogènes, ce qui résulterait en une induction de la prolifération de l'épithélium des conduits mammaires (7,8). Des études d'abolition du récepteur de l'EGF chez les souris ont démontré une réponse diminuée à l'EGF dans la glande mammaire. Les souris vierges de cette étude possèdent en effet un déficit dans la croissance des conduits et le branchement des lobes avéolaires. Ce type de déficit est habituellement présent seulement dans les cas où les niveaux d'oestrogènes circulants sont élevés et par conséquent où l'EGF est augmenté (18). Une autre étude appuie cette hypothèse; l'introduction de 17β -estradiol dans les tissus adipeux mammaires chez la souris ovariectomisée augmente les niveaux locaux d'EGF, qui lui-même semble induire les effets mitogéniques des oestrogènes (19).

1.2.3 Les tissus osseux

Un lien entre les hormones oestrogéniques et l'ostéoporose a été établi depuis plus de soixante ans (1941) chez les femmes ménopausées (8). Maintenant bien caractérisée, cette maladie provient d'un débalancement entre l'action des ostéoblastes et des ostéoclastes (8). Les os sont composés principalement de trois types de cellules : les ostéoblastes, qui participent à la synthèse de la matrice osseuse; les ostéoclastes, qui participent à sa résorption; et finalement les ostéocytes qui lient la masse osseuse (20).

Les hormones jouent un rôle dans le dimorphisme sexuel qui est atteint lors de la puberté. Les oestrogènes limitent alors la vitesse de croissance et jouent un rôle sur l'arrêt des plaques de croissance osseuses. Durant l'âge adulte chez la femme (avant la ménopause), le cycle menstruel influence l'activité des ostéoblastes indirectement. Pendant une grossesse, il y a maintien de la masse osseuse mais augmentation de la vitesse du cycle de synthèse et de résorption par les ostéoblastes et ostéoclastes, respectivement. Lors de la

ménopause, la diminution des œstrogènes circulants perturbe la balance entre l'activité de ces cellules osseuses. L'activité des ostéoclastes, n'étant plus ralentie par les œstrogènes, va augmenter et les ostéoblastes auront plus de difficulté à remplir les espaces résorbés. Ce manque d'œstrogènes va ainsi accroître les risques d'ostéoporose (20).

Les études chez les animaux et les cellules en culture appuient ces observations chez la femme. Bien qu'il reste encore des controverses et des imprécisions quant aux détails moléculaires du mode d'action des oestrogènes dans les tissus osseux, il semble y avoir un certain consensus. Les oestrogènes, agissant à la fois sur les ostéoclastes et les ostéoblastes, maintiennent le contrôle de la balance synthèse – résorption. L'activité de résorption des ostéoclastes serait influencée directement par les hormones oestrogéniques (7,20). Cependant, l'action directe des oestrogènes sur les ostéoblastes est plus discutable. Les oestrogènes influencent les cytokines sécrétées par les ostéoblastes (IGF1, IL1, IL6, etc...) (7,20), contrôlant ainsi leur activité de manière indirecte.

1.2.4 Le système nerveux

Plusieurs raisons sont évoquées pour justifier l'utilisation de thérapie de remplacement hormonal lors de la ménopause chez la femme. Certaines d'entre elles sont reliées aux effets des oestrogènes sur le système nerveux. Les oestrogènes pourraient prévenir les maladies dégénératives comme la maladie d'Alzheimer et celle de Parkinson (21,22). La détection des sites d'action précis des hormones dans le cerveau a démontré de nombreux sites de fixation (cortex, hypothalamus, région pré-optique, amygdale, cordon médullaire, épiphyse) (9). Cette distribution confère un rôle potentiel des hormones dans différentes sphères physiologiques, allant des fonctions reproductrices, de la mémoire et de l'apprentissage à la régulation hormonale (8,23). Il est également reconnu que les hormones sexuelles influencent la différenciation sexuée du système nerveux (8,9).

Comme dans beaucoup d'autres tissus lors de la maturité sexuelle, les oestrogènes jouent un rôle dans les fonctions de reproduction contrôlées par le système nerveux. Les niveaux d'oestrogènes circulants chez la femme régulent de manière positive et négative les

sécrétions de gonadotropines (FSH, LH) par l'hypothalamus. Ceci a pour rôle de contrôler la sécrétion des stéroïdes, la gamétogenèse et l'ovulation (8).

1.2.5 Le système cardiovasculaire

Il est reconnu depuis nombre d'années que les risques de maladies cardiovasculaires sont plus faibles chez les femmes pré-ménopausées que chez les hommes du même âge. Il est donc naturel de penser que les oestrogènes ont un rôle à jouer dans cette protection. Plusieurs études épidémiologiques ont en effet démontré l'implication des oestrogènes dans la protection cadiovasculaire (8,24-26).

Les oestrogènes protègent les femmes principalement de deux façons. Dans un premier temps, elles jouent un rôle positif sur le système cardiovasculaire en agissant directement sur les vaisseaux sanguins (8,24). Les oestrogènes modulent une vasodilatation en agissant sur les cellules musculaires lisses et l'épithélium des vaisseaux sanguins (24). Une panoplie de molécules est régulée par les oestrogènes [molécules d'adhésion, vasodilatateurs (NO), vasoconstricteurs (endothéline-1), anti-oxydants, chemoattractants] (8). Cette régulation dans le territoire veineux, complémentée par les effets des oestrogènes sur le cholestérol, contribue ainsi à une meilleure circulation sanguine.

Le deuxième rôle protecteur est indirect. Les oestrogènes participent au contrôle de l'hypercholestérolémie en venant diminuer les LDL circulant dans le sang et en augmentant la concentration de HDL (8,24). Des études chez des animaux ovariectomisés ont démontré qu'un anti-œstrogène total contrecarrait l'effet protecteur (27). La majeure partie de la protection provient de cet effet relié aux lipoprotéines (8).

1.3 Les maladies associées aux oestrogènes

Les oestrogènes ciblent donc des tissus très variés chez la femme et chez l'homme. Les actions variées des oestrogènes peuvent conduire à une protection contre certaines maladies (ostéoporose, Alzheimer, Parkinson, maladies cardiovasculaires...) ou à une

sensibilité accrue pour d'autres maladies (cancer des ovaires, de l'utérus, du sein, de la prostate, du colon). Par soucis de concision, seules les implications dans les cancers vont être abordées.

Bien que les oestrogènes protègent contre différents états pathologiques, une exposition prolongée peut au contraire induire différents types de cancer. Le type de cancer le plus fréquemment induit de cette manière chez la femme est le cancer du sein. Au Canada, une femme sur neuf en sera atteinte au cours de sa vie. Bien que les causes puissent être multiples, deux hypothèses se dégagent pour cette susceptibilité. L'exposition répétée des cellules mammaires à des cycles de prolifération, qui augmentent le nombre de divisions, augmente considérablement le risque d'erreur lors de la réplication ce qui peut conduire à des aberrations dans des voies cellulaires critiques comme l'apoptose et la prolifération. La deuxième hypothèse implique directement les métabolites actifs des oestrogènes qui peuvent induire des effets génotoxiques et provoquer ainsi des dommages à l'ADN (28).

Les récepteurs des oestrogènes sont exprimés dans deux tiers des tumeurs ovariennes. Ce cancer relié en grande majorité aux oestrogènes est le cinquième cancer en importance chez la femme aux États-Unis. Contrairement au cancer du sein, l'effet de diminution de la récurrence de ces tumeurs ovariennes grâce à l'anti-œstrogène tamoxifène n'est pas observable. Le cancer du colon est quant à lui caractérisé par une perte du récepteur des oestrogènes. Les études cliniques à grande échelle sur les thérapies de remplacement hormonal ont permis de mettre en lumière que les traitements aux oestrogènes sans progestines augmentaient considérablement les cas de cancers endométriaux. Finalement, chez les rongeurs, une exposition aux oestrogènes cause une hyperplasie de la prostate menant à un cancer. Il faut cependant noter que les études chez les humains sont plus controversées sur ce sujet, bien que ces tumeurs soient positives pour les récepteurs aux oestrogènes (28).

Les oestrogènes jouent un rôle crucial dans la reproduction humaine et dans l'homéostasie chez la femme. Les maladies associées à leur présence et/ou leur absence

sont diverses. Ces hormones induisent une panoplie de réponses dans les cellules des tissus cibles grâce à l'action de différents transmetteurs de signaux. Les différents récepteurs des oestrogènes permettent cette signalisation. Dans le chapitre suivant, les récepteurs, leurs rôles et signalisation vont être abordés.

Chapitre 2 : Récepteur des oestrogènes, médiateur de la signalisation cellulaire

ONE OF the challenging problems confronting biological scientists has been the manner in which hormones serve as regulators of biochemical processes in tissues of higher animals

E. V. Jensen and E. R DeSombre, 1973 (8,29)

Le groupe du Dr E.V. Jensen a passé plusieurs années à démontrer que les hormones stéroïdiennes lient des récepteurs intracellulaires. Le premier récepteur à être identifié fut celui liant le 17β -oestradiol (ou E2), qui est l'œstrogène le plus actif (8). Par la suite, les récepteurs de la testostérone, la progestérone, les glucocorticoïdes, l'hormone thyroïdienne, la vitamine D3 et les rétinoïdes ont été documentés, définissant ainsi la grande famille des récepteurs stéroïdiens (8).

Dans les années qui ont suivi la caractérisation du concept de récepteur liant les hormones, l'avènement de la biologie moléculaire a permis le clonage d'un récepteur intracellulaire des œstrogènes, ER α (30), suivi après plusieurs années de ER β (31,32). Dans les trois dernières décennies, l'élaboration d'anticorps sélectifs et de souris mutantes où l'expression des récepteurs a été abolie, la compréhension moléculaire des modes de transmission du signal des récepteurs, la génération de différents antagonistes sélectifs, l'étude des maladies associées aux hormones et à leurs récepteurs ainsi que plusieurs autres avancées technologiques et mécanistiques ont permis de broser un portrait complexe mais en pleine effervescence du mode d'action des hormones stéroïdiennes et particulièrement des œstrogènes.

Dans cette deuxième partie, les détails moléculaires de la transmission du signal par les oestrogènes vont être développés en portant une attention particulière aux connaissances actuelles sur les mécanismes de liaison des récepteurs à l'ADN et de leur régulation transcriptionnelle, aux voies de signalisation classiques et non classiques des récepteurs et à

la complexification du concept de récepteur des hormones stéroïdiennes par la description de nouveaux acteurs.

2.1 Les récepteurs nucléaires des oestrogènes

Les hormones oestrogéniques, possédant des propriétés hydrophobes, entrent dans la cellule et ont pour cibles des récepteurs nucléaires. Il existe deux types de récepteurs nucléaires classiques liant les oestrogènes: ER α et ER β . Le premier récepteur à être cloné fut ER α en 1985 (30). Par la suite, des études d'affinité de liaison de E2 sur des tissus animaux ont été réalisées dans le but de comprendre les effets physiologiques des oestrogènes. Certains effets ne semblaient pas être explicables par la seule activité de ER α , dont l'expression n'était pas détectée dans certains tissus cibles des oestrogènes. Il fallut attendre 1995 pour que plusieurs groupes identifient le deuxième récepteur nucléaire, ER β , permettant ainsi d'ouvrir la voie à une meilleure compréhension des effets médiés par les oestrogènes (31,32).

Ces deux récepteurs se comportent comme des facteurs de transcription dépendant de la liaison des oestrogènes. Ils possèdent donc plusieurs domaines fonctionnels pour établir des interactions spécifiques avec l'ADN et leurs différents corégulateurs. Voici les descriptions de ceux-ci dans le contexte de la signalisation par les oestrogènes.

2.1.1 Dualité de fonction : récepteur et facteur de transcription

Les récepteurs des oestrogènes font partie de la grande famille des récepteurs nucléaires, qui ont une organisation en différents domaines fonctionnels. Ils reçoivent le signal hormonal et le transmettent en agissant sur la transcription via leurs domaines de transactivations 1 et 2 (AF1 et AF2) ainsi que par le domaine de liaison à l'ADN (Fig. 2). Ce dernier se situe dans la région «C» et inclut une partie de la région «D» (Fig. 2). Cette liaison directe à l'ADN est assurée par deux doigts de zinc à quatre cystéines (Fig. 2). Ces derniers comportent deux régions critiques : la «P box», pour la sélectivité de la liaison à

l'ADN et la «D box» comme domaine de dimérisation entre les deux unités des récepteurs, ce qui contrôle la spécificité de l'espacement (33-35). Cette région d'interaction avec l'ADN étant la plus conservée entre les deux récepteurs (> 95%, acides aminés), la liaison à l'ADN est presque identique entre ER α et ER β (23,36).

Le deuxième domaine le plus conservé entre les récepteurs (60%) est celui liant l'hormone (LBD). Situé dans la région «E», il est chevauché par la région d'activation AF-2 (hélices 3,4,5 et 12), qui est dépendante de la liaison de l'hormone. La poche hydrophobe de liaison au ligand est composée des hélices 3 à 12. Cette dernière hélice est la clé de l'activation du récepteur. Lors de la liaison du 17 β -oestradiol, des changements conformationnels provoquent le recouvrement de la poche de liaison par l'hélice 12 (H12), formant avec les hélices 3,4,5 une surface de recrutement de plusieurs coactivateurs par la formation d'un sillon qui permet l'insertion d'une hélice α comportant un motif LXXLL retrouvé dans la plupart des coactivateurs des récepteurs nucléaires. Par contre, lors de la liaison d'un antagoniste, l'hélice 12 se retrouve plutôt dans un sillon hydrophobe formé par les hélices 3, 4 et 5, déstabilisant les interactions avec les coactivateurs et permettant le recrutement de répresseurs transcriptionnels. Bien sûr, le portrait est plus complexe lorsqu'on examine chaque ligand individuellement, mais en résumé, c'est la mobilité de l'hélice H12 dans le domaine AF-2, qui dicte les interactions avec les corégulateurs des récepteurs. Les mécanismes de liaison et l'affinité de ER α et ER β diffèrent très peu pour les hormones naturelles. Finalement, en plus de la liaison du ligand, la région E/F possède une région de dimérisation entre les monomères de ER et des signaux de translocation nucléaire (7).

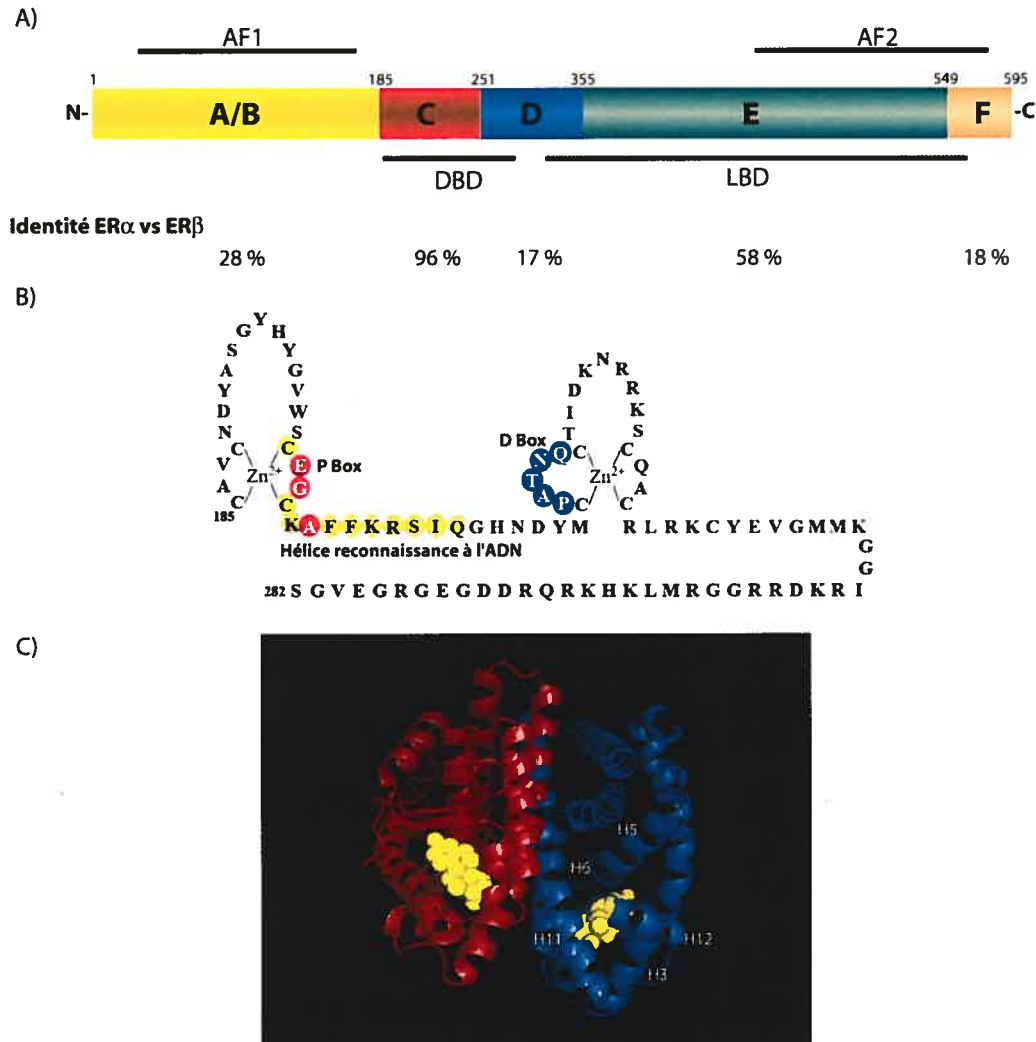


Figure 2 : Les récepteurs des oestrogènes

Les principales régions structurales des récepteurs des oestrogènes sont représentées. **A)** Schématisation de la séquence de ERα humain divisé en région A-F. Les deux principaux domaines fonctionnels du récepteur sont représentés, soit le domaine de la liaison à l'ADN (DBD) compris dans les régions CD et le domaine de liaison du ligand LBD débutant à la fin de la région D pour se terminer dans la région F. L'homologie entre ERα et ERβ pour les différentes régions est indiquée. Finalement on retrouve les régions d'interaction avec les corégulateurs AF1 (région A/B) et AF2 (région E/F). (23,36). **B)** Le domaine de liaison à l'ADN du récepteur ERα compris entre les acides aminés 185-282 dans les régions CD est élargi. On peut voir les deux doigts de zinc, en rouge, la P-BOX, en bleu, la D-BOX et en jaune l'hélice de reconnaissance à l'ADN. (33-35) **C)** Le domaine de liaison de l'hormone est représenté sous forme de dimère. Bien que les données cristallographiques aient été générées avec les acides aminés 301 à 553 du récepteur (DEF), seulement les résidus 305 à 548 sont structurés dans le cristal. On peut voir en jaune le 17β-estradiol lié dans sa poche de liaison hydrophobe, la position de la dernière hélice (H12) recouvrant le site de liaison est également visible. Finalement, l'interface de dimérisation (15% de la surface) comprend les hélices 8 et 11. (adapté de la séquence PDB de (37))

Finalement, la région la moins conservée (<16%) entre les deux récepteurs est la région A/B. Cette région est caractérisée par une fonction d'activation transcriptionnelle qui s'explique par la présence de différentes régions d'interaction avec des corégulateurs. Cette fonction de transactivation (AF-1) est indépendante du ligand. Il existe des différences majeures entre les domaines AF-1 des deux ER. Des essais réalisés dans différents types cellulaires ont démontré que le domaine AF-1 de ER α est beaucoup plus actif en transcription que celui de ER β . Cette différence réside peut-être dans l'identité des corégulateurs interagissant avec les deux récepteurs ou dans la force de ces interactions. Une autre différence dans les propriétés transcriptionnelles des deux récepteurs est due à l'activité de AF-1 lors de la liaison de différents ligands. L'anti-oestrogène tamoxifène (TAM), qui agit comme un agoniste partiel de manière dépendante des gènes et du type cellulaire, médie l'activation de la transcription grâce à certaines parties de la région AF-1 dans le récepteur ER α . Cependant lorsque les mêmes expériences sont faites avec le récepteur ER β , il n'y a aucune activation avec TAM. Cette activation est retrouvée lorsque des chimères de récepteurs pour la région A/B sont réalisées. Les 24 acides aminés (41-64) nécessaires à cette activation pour ER α ne se retrouvent pas dans le domaine AF1 de ER β , ce qui explique la différence d'activité transcriptionnelle face au TAM (7,38-40).

2.1.2 L'activation et les modifications post-transcriptionnelles

Comme tous les récepteurs et les facteurs de transcription, les récepteurs des oestrogènes peuvent être modifiés de façon post-transcriptionnelle. Ces modifications sont à la fois importantes pour l'activité de transactivation, la stabilité du récepteur, l'interaction avec des corégulateurs, la localisation dans la cellule, la liaison de l'hormone et la liaison à l'ADN. Il existe une variété de modifications post-transcriptionnelles sur des protéines telles que la phosphorylation, l'acétylation, la méthylation, la sumoylation, l'ubiquitination et la palmitoylation. La description exhaustive des différentes modifications démontrées sur ER pourrait être le sujet d'une thèse en soi, donc seuls seront décrits ici les principaux effets des modifications et leurs différents impacts sur l'activité du récepteur.

Lors de la liaison de l'hormone, les changements conformationnels du récepteur ont été démontrés par des modifications dans l'immunoréactivité du récepteur, dans des patrons de digestions par les protéases altérées et dans les surfaces hydrophobes exposées (41). Ces différents changements induits par la liaison de E2 permettent d'exposer des surfaces sensibles aux modifications et au recrutement des différentes protéines régulatrices. Ces dernières peuvent à la fois participer directement à la transcription médiée par ER, à la modification d'autres corégulateurs et/ou à la modification des récepteurs eux-mêmes. L'intervention de ces corégulateurs rend possible différentes modifications sur ER α , dont les plus importantes et/ou les mieux caractérisées sont la phosphorylation, l'acétylation, la sumoylation et l'ubiquitination.

La phosphorylation est sans contredit la modification la plus largement étudiée sur les récepteurs stéroïdiens. Elle est impliquée à la fois dans l'activité transcriptionnelle indépendante (AF-1) et dépendante (AF-2) des œstrogènes (41-45). Plusieurs sites de phosphorylation ont été caractérisés dans ER α , se situant tout au long du récepteur. Ils impliquent des sérines/thréonine kinases et des tyrosine kinases (42,45). Certaines phosphorylations de sérines dans le domaine AF-1 sont impliquées dans l'augmentation de l'activité transcriptionnelle indépendante (43,45) et dépendante (41-43,45) du ligand en permettant le recrutement de différents coactivateurs. Cette activité améliorée peut être due à la fois à une meilleure stabilité du récepteur, à un recrutement accru des coactivateurs, à une meilleure liaison à l'ADN ou à un changement de localisation du récepteur (41,42,45).

Le deuxième type de modification relativement bien caractérisé est l'acétylation du récepteur. Lors de la transcription par ER α , une variété de coactivateurs qui possèdent une activité acétyltransférase (HAT) sont recrutés (46). Dans un premier temps, ces modifications ont d'abord été reconnues pour être impliquées dans l'ouverture de la chromatine en venant déstabiliser sa structure compacte, ce qui favorise le recrutement des protéines servant à la transcription. Des études ont démontré que ces coactivateurs pouvaient à la fois agir sur les histones et sur le récepteur lui-même. Pour le moment, l'acétylation de quatre lysines de ER α (domaine D) est caractérisée, c'est-à-dire les résidus

266, 268, 302 et 303. La caractérisation des deux premières lysines découvertes (302, 303) indique que leur acétylation semble être impliquée dans une baisse de l'activité transcriptionnelle du récepteur (45,47). Cette observation est à priori en contradiction avec l'effet de l'acétylation des histones, qui augmente la transcription, et suggère un mécanisme de régulation en boucle de rétroaction du récepteur. Des études subséquentes ont démontré qu'une phosphorylation dans la même région pouvait prévenir l'acétylation de ces lysines (44). Cette phosphorylation est quant à elle activatrice de la transcription. Les deux modifications sont donc opposées. Cette acétylation inhibitrice pourrait avoir un rôle à jouer dans l'inactivation et la dégradation subséquente du récepteur. Quant à l'acétylation des lysines 266 et 268, elle est impliquée dans une meilleure activité transcriptionnelle de ER α en augmentant son affinité pour l'ERE (46). L'acétylation est donc une modification qui provoque des régulations complexes et diversifiées de ER α .

L'ubiquitination des ER est utilisée pour servir de ciblage vers la machinerie de dégradation des protéines impliquant le protéosome. La demi-vie de ER α est réduite à approximativement 3 heures en présence d'œstrogène. L'ubiquitination du récepteur, parallèlement à la transcription, sert de modulateur du signal de transcription induit par le ligand. Cette modification semble aller de pair avec une réduction de la mobilité intracellulaire du récepteur qui est par la suite dégradé (45,48). Elle est donc reliée à une inactivation du récepteur.

Finalement, la sumoylation, bien qu'appartenant à la même famille de modifications que l'ubiquitination, a un rôle positif dans la transcription médiée par les ER (49). Elle semble être présente seulement lorsqu'il y a liaison de E2 au récepteur (49). Contrairement à l'ubiquitination, elle ne joue pas un rôle de ciblage du récepteur vers le protéosome.

2.1.3 La liaison à l'ADN

Les récepteurs des oestrogènes ont la capacité de lier l'ADN via leurs doigts de zinc (Fig 2 et 3). Plusieurs études ont été réalisées pour déterminer quels sont les déterminants

de cette liaison et quelle en est la spécificité. Les récepteurs stéroïdiens lient l'ADN généralement avec deux unités de récepteurs (homo- ou hétéro- dimère), sur deux demi-sites consensus de cinq ou six paires de bases (pb). La spécificité d'interaction vient en grande partie de l'orientation que prennent ces demi-sites sur l'ADN et de l'espacement entre ceux-ci. Les séquences des sites eux-mêmes sont en effet très similaires entre les différents récepteurs nucléaires. Ces séquences peuvent être cependant organisées en palindromes, en répétitions directes et/ou en répétitions inverses, avec des espacements variables généralement de un à cinq nucléotides (36).

L'élément de réponse consensus pour les récepteurs des oestrogènes sur l'ADN (ERE) est illustré dans la figure 3A. Il est composé d'un palindrome de six paires de bases (pb), espacé par trois paires de bases d'identité variable. Bien que les positions extérieures des motifs soient plus permissives (purine/pyrimidine), il n'en demeure pas moins que cet ERE (Fig 3A) possède la plus haute affinité d'interaction pour les ER (33,35,36). Les autres récepteurs stéroïdiens (ex : récepteur glucocorticoïdes (GR)) lient le même genre de motif palindromique avec pour seules différences les deux bases centrales du demi-site. Des études ont démontré que la «P box » (Fig 2) est impliquée dans la discrimination entre les motifs reconnus par ER ou par les autres récepteurs stéroïdiens (35). Des études ont également caractérisé l'espacement (trois nucléotides) et l'orientation palindromique des demi-sites comme étant optimaux (33,35). ER α peut lier faiblement un demi-site répété directement avec une plus faible affinité. Il a également été établi qu'une petite partie de la région «D» (Fig 2) était impliquée dans une interaction avec l'ADN et aide la stabilisation de l'interaction, particulièrement dans les séquences présentant des variations avec le consensus (34). Des études cristallographiques ont permis la caractérisation en trois dimensions de la position des unités de récepteur sur une petite molécule d'ADN, et d'appuyer les études précédentes (50) (Fig.3 B,C). Dans la résolution des cristaux, on peut bien voir les acides aminés de la «P box» jouer un rôle majeur dans la reconnaissance des bases qui diffèrent entre les deux types de motifs reconnus par les récepteurs stéroïdiens.

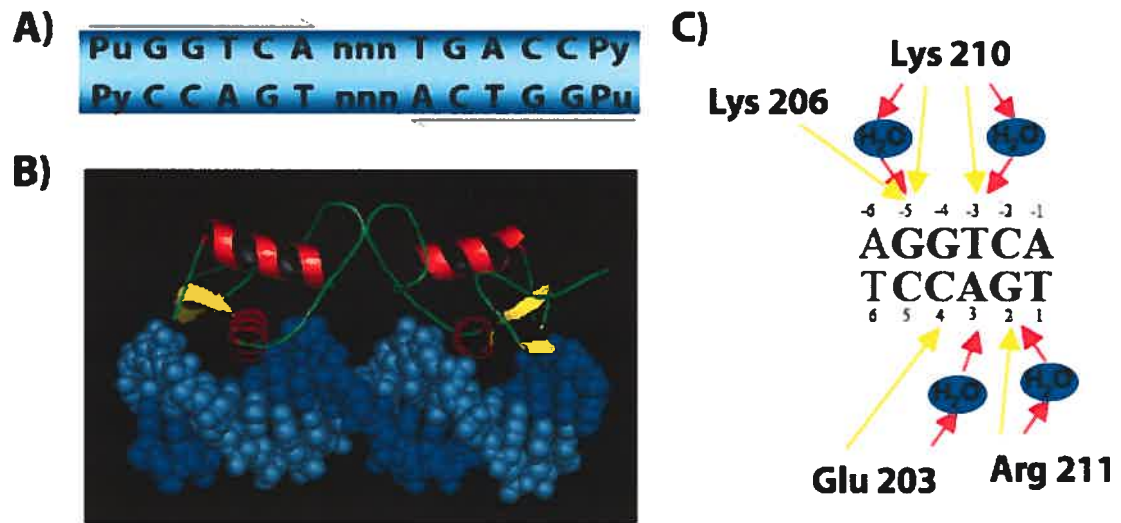


Figure 3 : Interaction de ER α avec l'ADN

A) Élément de réponse aux oestrogènes ERE. B) Modèle de liaison du DBD de ER α sur un ERE obtenue en utilisant un polypeptide correspondant à la région des doigts de zinc seulement. (Préparé à partir de la structure de la banque de données PBD correspondant à la référence (50)) C) Interactions directes et indirectes des différents acides aminés du DBD avec l'ADN (36). (figure bâtie à l'aide des données cristallographiques (50) et de la référence (36))

Certaines études se sont penchées sur la différence potentielle de liaison qui existe entre ER α et ER β . Le site de liaison consensus est le même et il y a de très petites variations d'affinité entre ER α et ER β (51). Les différences entre l'affinité de ces récepteurs envers le ERE *in vitro* résident dans leur affinité envers les séquences variantes du consensus. Des études plus poussées d'immunoprécipitation de la chromatine devront être réalisées pour discriminer si les récepteurs agissent différemment dans un contexte de chromatine structurée *in vivo*.

2.1.4 La distribution tissulaire

Dans le premier chapitre, les effets des oestrogènes sur les différents tissus ont été abordés. Ces effets variés sont médiés par les récepteurs ER α et ER β qui, une fois activés par l'oestrogène et liés à l'ADN, régulent la transcription d'une panoplie de gènes différents

selon les tissus. Dans l'organisme, les récepteurs sont retrouvés parfois dans les mêmes tissus, avec un rapport $ER\alpha/ER\beta$ différent. Ces différences d'expression permettent une large modulation de la réponse à un même signal par les oestrogènes.

Le récepteur $ER\alpha$ est le seul ER exprimé dans les ovaires, le foie, les reins et l'hypophyse (8). Par contre, dans les glandes mammaires, les os, l'utérus, les ovaires, les poumons, le système nerveux, le système cardiaque, les testicules, la prostate et l'épididyme, les deux récepteurs sont présents avec des rapports de concentration variés (7,8). Cette liste établie chez le rat et la souris dépend cependant des souches de souris testées (www.nursa.org). Les techniques utilisées pour la détection influencent les résultats des patrons d'expression des récepteurs. L'âge des animaux testés joue également un rôle important dans l'expression des récepteurs. Cependant, bien que les tissus ne soient pas tous identiques d'un organisme à l'autre, il est important de retenir que les récepteurs sont exprimés de façons différentes dans une panoplie de tissus cibles. Ces variations brossent un tableau complexe du rôle des oestrogènes et de sa signalisation. Les gènes dont la transcription est régulée par les récepteurs pourront donc être affectés de manière différente dans les tissus, dépendamment de quels récepteurs y sont exprimés, quels sont les partenaires de corégulation présents, l'accessibilité de ces gènes aux récepteurs, etc...

2.1.5 Les récepteurs orphelins reliés aux oestrogènes (ERR)

Après le clonage de $ER\alpha$ et d'autres récepteurs nucléaires (récepteurs aux hormones thyroïdiennes, de la vitamine D3) particulièrement proches des ER, plusieurs criblages ont été réalisés dans le but de découvrir d'autres récepteurs nucléaires potentiels sur la base de la conservation des domaines de liaison à l'ADN. En 1988, un groupe a identifié des récepteurs nucléaires orphelins possédant un domaine de liaison à l'ADN similaire à $ER\alpha$. Ces deux récepteurs appelés «orphelins» car leur hormone régulatrice n'est pas connue, ont été nommés $ERR\alpha$ et $ERR\beta$ pour récepteur relié aux récepteurs des oestrogènes. Un troisième membre de cette famille de récepteurs orphelins, $ERR\gamma$ est venu s'ajouter en 1999. (52)

Les ERR possèdent une homologie assez grande (60%) avec ER α au niveau du domaine de liaison à l'ADN. Cette ressemblance permet aux ERR de lier l'ADN en monomère sur un ERE. De plus, ils sont capables de lier l'ADN sur un demi-ERE précédé des nucléotides TCA (le SFRE). Il a été démontré *in vitro* que ER α (mais pas ER β) est également capable de lier les SFRE. Certains gènes régulés par les oestrogènes via les ER, tels que la lactoférine, l'ostéopontine et le gène pS2 (TFF1), sont également régulés par les ERR. La liaison d'éléments d'ADN similaires entre les deux types de récepteurs complexifie grandement l'étude de la régulation par les oestrogènes (7,52).

Bien que leur interaction à l'ADN soit similaire, le domaine de liaison du ligand de ER α et de ERR présente seulement 35% d'homologie. Par conséquent, les oestrogènes naturels ne peuvent interagir avec les ERR, qui sont constitutivement actifs et peuvent intervenir sur différents gènes avec ER α à la fois de façon synergique ou opposée. Certains composés synthétiques, comme le DES et l'antagoniste de ER Tamoxifène, peuvent réprimer l'activité constitutive des ERR. Pour le moment, aucun ligand naturel n'est connu pour les ERR (7,52).

Ce croisement des voies de signalisation des récepteurs classiques des oestrogènes et des récepteurs orphelins peut donner lieu à une complexification de la régulation par les oestrogènes, puisqu'en plus de lier l'ADN de la même façon, les ERR régulent l'aromatase, enzyme responsable de la conversion des androgènes en oestrogènes.

2.2 La signalisation des œstrogènes

Les oestrogènes, de petites molécules hydrophobes, passent au travers des membranes cellulaires et se lient aux récepteurs nucléaires. Par la suite, les récepteurs effectuent dans le noyau la réponse transcriptionnelle adéquate par plusieurs mécanismes. En plus d'une liaison directe à l'ADN, mécanisme dit de signalisation classique, les récepteurs liant les oestrogènes peuvent participer à la transcription de gènes en collaborant

avec d'autres facteurs de transcription, directement sur l'ADN ou en tant que corégulateur. Finalement, les oestrogènes possèdent une voie de signalisation très rapide par un (des) récepteur(s) membranaire(s), dont l'identité est encore controversée. Dans les prochaines sections, les mécanismes de transmission du signal hormonal de façon classique seront abordés, ainsi que les nouvelles avenues de signalisation des oestrogènes.

2.2.1 Signalisation nucléaire classique

En absence d'oestrogènes, les récepteurs ER α et ER β se retrouvent à la fois dans le cytoplasme et dans le noyau. Ils sont couplés à des protéines chaperonnes (HSP90) qui les maintiennent dans un état inactif. Lors de la liaison de l'oestrogène au récepteur, il y a dissociation des protéines chaperonnes et changement conformationnel permettant la dimérisation des récepteurs et l'association subséquente avec les ERE accessibles sur l'ADN (Fig. 4A). L'association du récepteur sur l'ADN permet de médier la transcription d'un gène donné. Cette transactivation peut se produire seulement avec l'aide de plusieurs coactivateurs.

2.2.1.1 Les cofacteurs de la signalisation nucléaire

Plusieurs classes de cofacteurs interagissent directement ou via un intermédiaire avec le dimère de ER avec E2 lié sur l'ADN afin de permettre une réponse adéquate et complète du signal. Les différents coactivateurs seront recrutés différenciellement avec ER selon le gène et la cellule étudiés. Les principaux complexes pouvant participer à la transcription avec ER α sont : le complexe de remodelage de la chromatine SWI/SNF, les complexes à activité HAT (p160 et CBP/p300), les protéines avec une activité méthyltransférase (CARM1, PRMT1), et finalement le complexe médiateur TRAP/DRIP.

Selon une étude exhaustive du recrutement des différents complexes protéiques participant à la régulation de la transcription par ER α du gène pS2/TFF1, le premier coactivateur venant aider ER α sur le promoteur est Brg1 (53), une sous-unité catalytique du complexe des SWI/SNF. Ce dernier participe à la régulation de la transcription à l'aide de

l'activité hélicase dépendante de l'ATP, possédée par certaines sous-unités, en permettant un remodelage des nucléosomes (54). L'interaction hormono-dépendante de ce complexe protéique avec ER α a été découverte en premier lieu par des expériences de doubles hybrides chez la levure (55). Par la suite, plusieurs études ont permis d'étudier le rôle de ce complexe de cofacteurs avec ER α et les autres récepteurs nucléaires (53,56,57). En plus de jouer sur la structure du nucléosome, SWI/SNF (via BAF57) interagit avec des membres de la famille des p160, une autre classe de cofacteurs de ER α (57).

Toujours selon l'étude sur le gène TFF1, les p160, (SRC1, SRC2, SRC3) (58-63) possédant presque tous (SRC1, SRC3) (62,64) une activité histone acétyltransférase (HAT), arrivent à la suite du complexe SWI/SNF sur le promoteur (53). En présence d'œstrogène, ils interagissent avec ER α , via la région d'interaction formée par les hélices H3, H5 et H12 du LBD (AF-2) par un motif d'interaction LXXLL (65). Ces p160 peuvent recruter d'autres complexes protéiques, dont le complexe d'intégration CBP/p300, qui possède également une activité HAT(62,66) et les méthylases des histones (CARM1/PRMT1) (67), pour permettre une synergie du signal de transcription (68).

Ces différents complexes protéiques modifieront l'état de la chromatine et sa permissivité pour la transcription de façon à permettre l'arrivée des facteurs généraux de la transcription via le complexe médiateur qui sert de pont entre le récepteur/cofacteurs et les facteurs généraux de la transcription et la RNA polymérase II (53,69).

2.2.1.2 Recrutement ordonné des complexes impliqués dans la signalisation classique

Plusieurs études ont déterminé que ces différentes protéines ne se retrouvent pas toutes au même moment avec le récepteur sur l'ADN, mais elles sont recrutées de façon ordonnée. Pour le gène TFF1, les coactivateurs SRC/p160 et CBP/p300 sont mutuellement exclusifs. Ces recrutements sont ordonnés et séquentiels mais sont également variables selon le promoteur et la cellule (36,70).

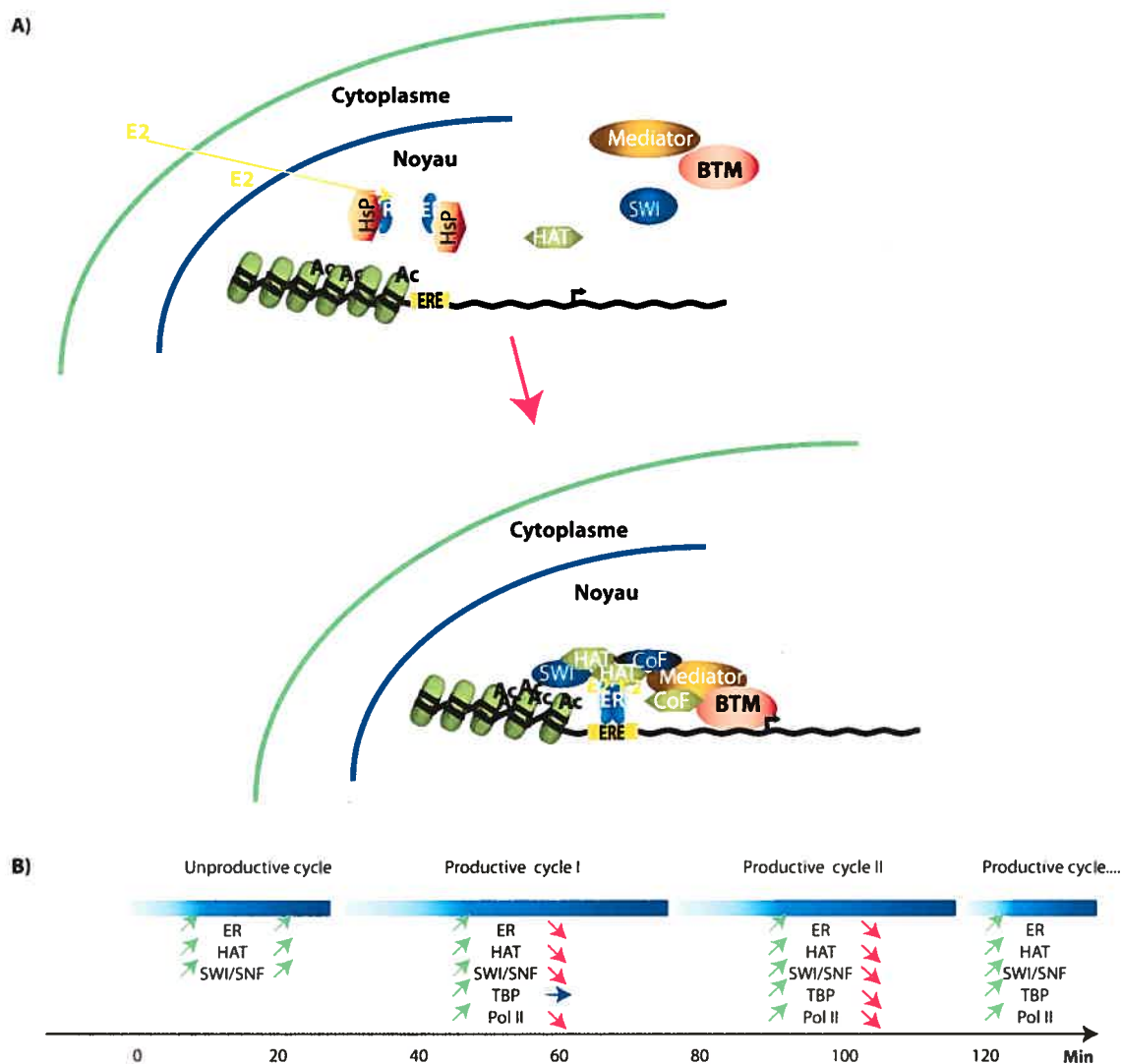


Figure 4 : Modèle de signalisation dynamique et directe de ER α . A) Signalisation via le mode classique des récepteurs des oestrogènes. Lors de l'arrivée de l'oestrogène, les protéines de choc thermique HSP se dissocient, il y a changement conformationnel du récepteur, dimérisation, liaison du récepteur sur un site ERE et recrutement des différents coactivateurs. B) Les coactivateurs sont recrutés de manière dynamique sur le promoteur de TFF1. Voici un sommaire des différentes interactions avec ER α lors de la transcription. En vert : augmentation, en rouge diminution, en bleu recrutement stable (adaptation de Métivier et al (53,70)).

Des études d'immunoprécipitation de la chromatine ont permis de déterminer que le recrutement des différents complexes n'est pas un évènement statique dans le temps mais très dynamique. Il a été démontré que l'association de ER α et de l'ARN polymérase II est périodique (Fig. 4B). L'utilisation d'un agent chimique (α -amanitine) qui permet la

synchronisation de la transcription en bloquant l'ARN polymérase II a permis de démontrer la présence d'un cycle d'association non productif du récepteur en absence de ligand. Après l'ajout d'E2, un premier cycle d'association non productif permet de préparer le promoteur aux cycles d'association de ER α et des différents coactivateurs qui permettent la transcription (les cycles productifs I et II sont répétés en alternance) (Fig 4). La présence de la polymérase II activée concorde avec une diminution de ER α sur le promoteur. Cette période réfractaire au récepteur permet le recrutement de différents complexes corépresseurs (HDAC) entre les cycles de transcription de façon à diminuer l'accessibilité pour le ERE par ER α . Le recrutement de protéines appartenant au complexe du protéasome 20S et la présence de protéines ubiquitinées à proximité du promoteur laissent également entrevoir la possibilité que ER α soit dégradé simultanément durant ces cycles «d'inactivation» de la transcription (53,70-72).

Ces différentes études ont démontré que la transcription classique par ER α est un processus dynamique et cyclique, impliquant différentes classes de protéines. La majorité de ces études, réalisées sur un nombre restreint de promoteurs, se sont concentrées exclusivement sur les gènes dont la transcription est activée par ER α . Plusieurs autres études devront être réalisées sur différentes classes de gènes pour établir un schéma global de la transcription médiée par les ER.

2.2.2 Signalisation nucléaire alternative

Les récepteurs des oestrogènes sont connus pour lier l'ADN via leur ERE dans la majorité des situations. Cependant, il existe des modes alternatifs de régulation de la transcription. Le récepteur peut lier l'ADN en combinaison avec un autre récepteur nucléaire et/ou facteur de transcription. Cette liaison peut être directe via des éléments de réponse combinés ou indirect, agissant alors comme coactivateur activé par l'oestrogène. Il existe donc deux grandes catégories de liaison alternative : liaison nécessitant l'interaction directe entre ER α et l'ADN, et le recrutement indirect du récepteur à l'ADN via un autre facteur de transcription (36,73-75).

Le récepteur ER α , contrairement à ER β , est capable de lier l'ADN à un demi-ERE (SFRE). Cet élément de réponse peut être seul sur le promoteur ou souvent combiné avec un site de liaison pour un autre facteur. Ce cas est bien documenté avec le facteur de transcription SP-1. Plusieurs études ont démontré que des sites combinés fonctionnels existent. Les distances entre le demi-ERE et la région riche en GC (site SP-1) ainsi que leur orientation sont variables d'un gène à un autre. Un autre exemple est l'interaction de ER α avec NF- κ B, où le DBD de ER α est requis (36,74-76).

L'interaction indirecte de ER α avec d'autres facteurs de transcription sans la nécessité du domaine de liaison à l'ADN a été démontrée pour quelques gènes. ER α est capable de réguler l'expression de différents gènes en agissant comme coactivateur de SP-1 et d'AP-1 en particulier. Des études à grande échelle ont d'ailleurs été faites avec AP-1 pour identifier des gènes régulés par les oestrogènes de cette façon particulière (36,74,75,77).

Ces modes de signalisation alternatifs ne semblent pas changer les autres intervenants recrutés dans l'activation de la transcription. Cependant, peu d'études ont été réalisées dans le but d'analyser le recrutement des différentes composantes pouvant réguler la transcription via ER α dans un mode de transmission du signal alternatif.

2.2.3 Signalisation membranaire : mode d'action rapide des oestrogènes

Finalement, la signalisation des oestrogènes passe aussi par une composante rapide, souvent nommée voie «non-génomique d'activation» par opposition à la signalisation classique et/ou alternative qui se passe au niveau génomique (directement sur l'ADN). L'étude de cette signalisation non génomique est plus controversée. Elle passe par l'intervention de protéines membranaires et/ou associées à la membrane. La controverse provient de l'identification des effecteurs du signal oestrogénique: est-ce la présence d'un nouveau récepteur des oestrogènes membranaire et/ou un recrutement des ER à la membrane.

Les premiers effets rapides des stéroïdes ont été caractérisés en 1942 avec l'étude de la progestérone. Il fallut attendre jusqu'en 1975 pour la caractérisation de l'effet rapide des oestrogènes. Cette voie d'activation est mise en marche seulement de quelques secondes à quelques minutes après un stimulus. Comme les autres modes de signalisation membranaire, il y a stimulation des voies MAPK, PI3K/AKT, cAMP, voies calciques et induction du flux ionique (78,79).

Différentes études ont démontré que ER α et ER β pouvaient être recrutés à la membrane plasmique et transmettre ces signalisations rapides. Environ 5% des récepteurs se retrouvent à la membrane via leur interaction avec différentes protéines (les Shc) et forment un complexe dans les cavéoles avec les récepteurs membranaires IGF-1R et/ou EGFR. La majeure partie du recrutement de ER α à la membrane se fait avec l'aide de la protéine cavéoline-1. De plus, des sites de palmitoylation ont été caractérisés sur le récepteur, permettant un ancrage à la membrane. La présence de ER α dans les cavéoles permet au récepteur d'activer des voies de signalisation conduisant à la synthèse de eNOS. L'existence de cette signalisation, impliquant directement les récepteurs classiques recrutés à la membrane, a été appuyée par la génération de souris transgéniques où les gènes pour ER α et pour ER β ont été détruits. Ces souris présentent une absence de signalisation rapide dans les cellules épithéliales (78-81).

Ces dernières années, un récepteur orphelin couplé aux protéines-G, GPR30, a été caractérisé comme étant le médiateur potentiel des voies de signalisation non génomiques des oestrogènes. Les groupes travaillant sur cette hypothèse ont démontré que GPR30 peut lier E2 et activer les voies rapides de signalisation par les oestrogènes. De plus, GPR30 peut lier différents antagonistes de ER et toujours activer ces voies de signalisation rapides, contrairement à ER α qui est inhibé par ces molécules. Ces effets de GPR30 ont été démontrés dans des cellules exprimant aussi ER α (MCF7) et dans des cellules caractérisées pour être ER α - et ER β -négative (SKBR3). Récemment, un groupe a cependant mis un bémol à ces études en démontrant que dans leurs conditions expérimentales, l'activation

des voies rapides nécessite la présence de ER α et qu'il n'y a aucune signalisation par les oestrogènes par GPR30 dans les SKBR3 (78,79,82-84).

Bien qu'il y ait des controverses entourant l'identité du ou des récepteurs responsables de l'activité rapide à la membrane, il n'en demeure pas moins que les oestrogènes possèdent une signalisation rapide. Il est facile d'imaginer que les ER soient recrutés dans les cavéoles à la membrane dans les cellules les exprimant. Cependant, le mécanisme serait différent dans des cellules ne possédant pas ou peu de ER, laissant ainsi la place à des récepteurs comme GPR30. Les deux mécanismes n'étant pas exclusifs, il est possible que les deux écoles de pensées aient raison quant à l'identité de l'effecteur.

Les récepteurs ER α et ER β médient la signalisation des oestrogènes en intervenant directement ou indirectement sur la transcription des gènes cibles. Leurs actions peuvent être rapides et/ou plus lentes et prolongées. Les régulations et les effets induits par la liaison de E2 sont très diversifiés et complexes. Cependant, bien que le mode de transmission classique du signal soit bien caractérisé, on connaît encore peu d'éléments de réponse naturels et de gènes régulés de façon directe par les oestrogènes.

Chapitre 3 : Chromatine et transcription des gènes

Les principes généraux de la transcription étaient établis plusieurs dizaines d'années avant de connaître la séquence du génome humain. Bien qu'il y ait plus de 3 milliards de nucléotides dans le génome humain, les territoires géniques représentent seulement 5% de ce génome. Le reste de l'ADN possède différentes fonctions, dont entre autre la régulation de la transcription des gènes.

Lors de la transmission d'un signal à la cellule, la durée et l'amplitude de cette signalisation est contrôlées de plusieurs façons. La transcription des gènes est une des étapes importantes de la régulation de la signalisation dans une cellule eucaryote. Plusieurs intervenants protéiques collaborent avec les facteurs de transcription et les récepteurs nucléaires pour médier la réponse transcriptionnelle aux différents stimuli cellulaires. Outre la régulation directe de l'activité de ces différents intervenants, un autre niveau de régulation de la transcription existe et implique les composantes structurales des chromosomes.

Dans ce dernier chapitre, le concept de promoteur et les mécanismes généraux de la transcription seront décrits brièvement. Par la suite, l'organisation de la chromatine et son impact sur la transcription des gènes par les récepteurs nucléaires seront abordés. Finalement, l'émergence de la régulation de la transcription sur de grandes distances sera présentée.

3.1 L'organisation spatio-temporelle de la chromatine

Les chromosomes sont la forme tridimensionnelle la plus condensée de l'hélice d'ADN. La double hélice, passant de 2 nm à 1400 nm est enroulée, de manière très compacte, en association avec diverses protéines. Lors de la transcription d'un gène, les

différents coactivateurs permettent le relâchement de la structure de façon à donner accès à la machinerie de transcription. Dans les sections suivantes, la structure linéaire des promoteurs et la structure tridimensionnelle de la chromatine sera révisée, ainsi que les mouvements dynamiques de la chromatine.

3.1.1 Les composantes des promoteurs

Les gènes sont transcrits par la machinerie transcriptionnelle de base qui elle-même est recrutée à l'aide des facteurs de transcription et/ou des récepteurs nucléaires. Plusieurs études se sont penchées sur la description du promoteur minimal médiant la transcription basale. Bien qu'il n'existe pas de promoteur universel pour tous les eucaryotes, il y a plusieurs structures et motifs identifiés comme pouvant faire partie du promoteur minimal (85,86).

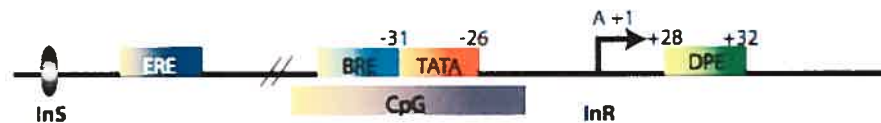


Figure 5 : Le promoteur minimal eucaryote. Le promoteur minimal correspond aux séquences -35 à +35 du site d'initiation de la transcription (Inr). Il existe plusieurs variations aux différentes séquences qui peuvent être retrouvées sur le promoteur minimal. La boîte TATA et l'élément DPE peuvent servir au recrutement du complexe TFIID, qui fait partie des facteurs généraux de la transcription. Dans d'autres cas, particulièrement dans les promoteurs riches en îlot CpG, il peut y avoir recrutement de TFIIB (un autre complexe membre des facteurs généraux de la transcription) sur un élément BRE. Finalement, les promoteurs possèdent, à différentes distances, des éléments pour les différents facteurs de transcription spécifiques à un gène donné, ici représenté par un ERE.

Le promoteur minimal est caractérisé par plusieurs sites de liaison pour différents complexes protéiques faisant partie de la machinerie transcriptionnelle de base. Ces différents sites de régulation et de liaison composant le promoteur minimal ont été identifiés à l'aide de différentes techniques biochimiques comme : la technique de *DNAse footprinting* (technique qui permet de voir l'occupation d'une région d'ADN par une protéine), des essais de gène rapporteur (technique qui permet d'étudier l'importance d'une région sur la transcription en utilisant des vecteurs d'ADN contenant un promoteur qui

régule la transcription d'une protéine luminescente) et la technique d'extension de la région 5' de l'ARN par amorce (5' RACE, technique permettant de déterminer le site d'initiation de la transcription d'un gène). La première région à avoir été caractérisée est la boîte TATA, motif d'ADN qui est similaire à la boîte Pribnow chez les procaryotes. Cette boîte est le site de liaison de TBP, une protéine composante du complexe TFIID, un facteur général de la transcription pour la polymérase II. Elle est située en amont du site d'initiation de la transcription (-31 pb à -26 pb, Fig 5). Dans les promoteurs ne possédant pas la boîte TATA, une autre région, située en aval du site d'initiation, peut servir à la liaison de TFIID : la région DPE. La région DPE est toujours localisée à la même distance du site d'initiation, soit +28 pb à +32 pb. Cet élément sert à positionner adéquatement le complexe de pré-initiation de la transcription (Fig 5). La dernière région caractérisée du promoteur minimal est une région riche en nucléotide GC. Cette région BRE, située immédiatement en amont de la boîte TATA, est le site de liaison de TFIIB, facteur général de la transcription qui peut activer et/ou réprimer la transcription dépendant du contexte des gènes et des organismes étudiés (85,86).

Le site d'initiation de la transcription (Inr) est présent dans les promoteurs contenant des boîtes TATA et/ou DPE. Le complexe TFIID et la RNA polymérase II peuvent lier la région Inr à différents moments de la transcription. Habituellement, la transcription commence au nucléotide A⁺¹. Cependant, certains gènes sont contrôlés par des facteurs de transcription liant des îlots CpG plutôt que d'être régulés par une boîte TATA et/ou DPE, la définition du nucléotide de départ est donc plus incertaine. Certaines études démontrent que près de la moitié des gènes humains étudiés seraient transcrits par un tel mécanisme. Ces îlots CpG sont caractérisés non pas par un site d'initiation précis (A⁺¹), mais plutôt par une transcription faible sur environ 100 pb (85-87).

Le rôle du promoteur minimal est d'assurer une transcription basale des gènes. Un gène nécessite beaucoup plus que les facteurs généraux de la transcription pour être transcrit et régulé adéquatement. C'est pourquoi on retrouve, à différents endroits sur le promoteur, des sites de fixation pour des facteurs de transcription spécifiques. Ces derniers

peuvent agir comme activateurs transcriptionnels et/ou comme inhibiteurs de la transcription basale. Pour certains facteurs de transcription comme $ER\alpha$, il y a recrutement du facteur au promoteur lors de l'arrivée du signal (dans ce cas-ci liaison d'oestrogènes au récepteur). Pour d'autres facteurs de transcription comme les récepteurs de la famille des rétinoïdes, ils peuvent être associés en absence de signalisation et médier une répression du gène. Cependant, lors de l'arrivée d'un signal positif de transcription, ces mêmes facteurs recrutent plutôt des protéines servant à l'augmentation de la transcription du gène. Le tableau peut être encore plus complexe car chaque promoteur peut contenir plusieurs sites activateurs et répresseurs différents et ces différents facteurs peuvent interagir entre eux de façon à promouvoir une régulation fine du gène. Ces sites de fixations peuvent également se situer sur de grandes distances par rapport au site d'initiation de la transcription, à la fois en amont et/ou en aval, et sont alors appelés sites «enhancers» ou «silencers». Ces derniers regroupent en général plusieurs sites de fixation pour différents facteurs de transcription. Finalement, il existe un dernier niveau de régulation : les régions frontières (Ins). Ces dernières servent à la fois de limites entre les domaines transcriptionnels et de barrières à la propagation de la structure de la chromatine régulant ainsi la propagation du signal des facteurs transcriptionnels (85-88).

3.1.2 Structure tridimensionnelle et dynamique de la chromatine

L'ADN contenu dans chaque cellule, lorsqu'il est déroulé, totalise environ 2 m de long. Il est donc nécessaire pour les cellules de posséder une façon efficace d'empaquetage de la molécule, car le noyau ne fait approximativement que 6 μm de diamètre.

Le premier niveau de compaction est obtenu par enroulement de l'ADN autour de protéines abondantes et hautement conservées dans la cellule : les histones. Des recherches de biochimie, de microscopie et de cristallographie ont permis d'établir que les histones (H2A, H2B, H3, H4) enroulent 146 pb en 1,75 tours formant ainsi, à la manière d'un collier de perles, une chaîne de 10 nm (89-91). Par la suite, les nucléosomes (histones + ADN) sont super-enroulés pour former une fibre de 30 nm. Finalement, le tout est à nouveau

enroulé en différentes boucles et super-hélices pour former un chromosome d'une largeur d'environ 1400 nm (<http://www.ncbi.nlm.nih.gov/Literature/> livre «*Molecular biology of the cell*»). Lors de la transcription, certains mécanismes de décondensation permettent une structure dynamique des fibres de chromatine. Les protéines recrutées par les récepteurs nucléaires permettent l'accessibilité à l'ADN de la machinerie transcriptionnelle de base en modifiant les différentes protéines structurales de la chromatine.

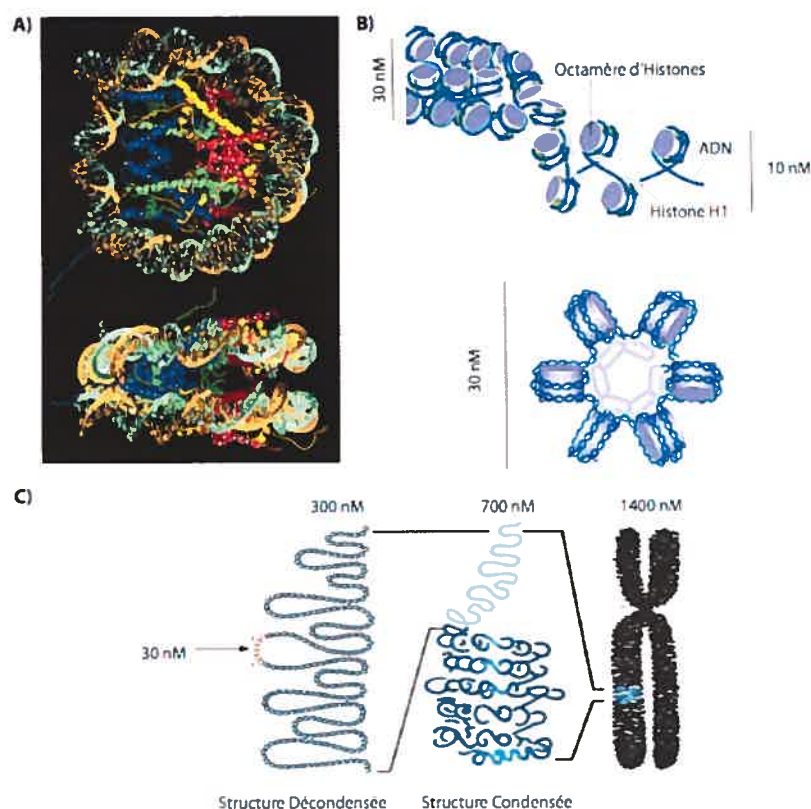


Figure 6 : Structure de la chromatine. A) Nucléosome formé d'un octamère d'histone, H2A, H2B, 2XH3, 2XH4 et 1,75 tours de la double hélice d'ADN (image tirées du livre Lodish et al, 2000, <http://www.ncbi.nlm.nih.gov/Literature/>). B) La chaîne de 10 nm formée d'une suite de nucléosomes, enroulée en fibres de 30 nm qui contient 6 nucléosomes. Les histones H1 sont utilisées pour maintenir une structure compactée. C) Par la suite, les fibres de nucléosomes sont super-enroulées dans des structures complexes. (Figure réalisé avec différentes images tirées du site <http://www.ncbi.nlm.nih.gov/Literature/>)

Les histones sont les protéines structurales à la base des nucléosomes. Elles étaient reconnues à l'origine pour être une composante statique de la chromatine (92). Cependant, plusieurs années de recherches ont démontré qu'elles jouent un rôle dynamique très important dans la régulation de la transcription par les récepteurs nucléaires et les autres classes de facteurs de transcription. Ce sont des protéines qui peuvent être modifiées de manière post-transcriptionnelle. Ces modifications (acétylation, méthylation, phosphorylation, ubiquitination, sumoylation, ribosylation et glycosylation) forment un code servant à réguler à la fois la transcription, la recombinaison, la réparation et la réplication de l'ADN (89,90,92). Les modifications dictent le type de protéines qui interagissent avec les nucléosomes grâce à leur domaine de liaison protéique. La structure super-enroulée n'est pas statique, quelles que soient les modifications effectuées aux histones. Il a été démontré que les nucléosomes oscillent entre un état compact (pendant 250 milli-secondes) et un état plus lâche (10-50 milli-secondes) sans l'intervention d'autres protéines et/ou de modifications post-transcriptionnelles (91,93).

Le nombre de gènes étant modeste comparativement à la taille du génome, de très grandes parties d'ADN ne servent pas à coder des protéines et des petits ARN de toutes sortes. Ces différentes régions «inutilisées», tout comme les régions codantes inactives, peuvent être inactivées par sur-enroulement de la chromatine et la formation d'une région compacte, l'hétérochromatine. Les histones forment cette structure compacte en étant méthylées à certains endroits précis sur leur extrémité, ce qui peut entraîner la méthylation de l'ADN et la subdivision de la chromatine en deux états. L'hétérochromatine est méthylée, très condensée, et transcriptionnellement peu active (94). Au contraire, l'euchromatine est peu méthylée, moins condensée et transcriptionnellement active. La majorité de la transcription d'une cellule va se produire dans l'euchromatine (93,95). Cette différenciation des régions transcriptionnelles actives et inactives a été observée dans les chromosomes polythènes de drosophiles. L'observation de cellules provenant de différentes larves de drosophiles (contenant des réarrangements chromosomiques différents) a permis de conclure que lorsqu'un gène donné se retrouvait dans une région «compacte ou noire» du chromosome on n'observait aucune transcription de ce gène dans la cellule. À l'opposé,

lorsque le gène se retrouvait dans des régions «décondensées ou pâles», il y avait transcription du gène (96,97). Ces observations ont été le prélude à la compréhension du rôle de la structure de la chromatine sur la transcription des gènes.

L'état plus lâche des nucléosomes, induit ou spontané, permet de découvrir sur de courtes périodes de temps certains sites de liaison des facteurs de transcription. Lors de l'arrivée d'un signal hormonal, le récepteur ER α (ou les différents facteurs de transcription) a donc la possibilité de lier les ERE accessibles. Par la suite, les coactivateurs (ex : p160) peuvent acétyler les régions N-terminales des histones. Cette acétylation vient déstabiliser les interactions entre l'ADN et les histones en neutralisant les charges positives des résidus lysines des protéines. Il s'en suit une déstabilisation des nucléosomes. Le recrutement du complexe SWI/SNF par le récepteur permet de déstabiliser les interactions des histones via une activité ATPase. L'action de SWI/SNF produit un déplacement des nucléosomes sur le brin d'ADN. Le résultat du recrutement de ces différents complexes protéines est que les histones seront acétylées et que la chromatine sera plus accessible aux autres corégulateurs de la transcription. Cette acétylation des nucléosomes est d'ailleurs propagée par les différents corégulateurs, permettant de rendre accessible le promoteur minimal à la machinerie transcriptionnelle de base en utilisant les protéines du médiateur comme pont protéique entre le site de liaison du facteur de transcription et le site d'initiation (91,92,98).

Plusieurs autres modifications des histones sont induites par des corégulateurs de la transcription. Ces modifications, qui ont lieu sur les queues terminales des histones, forment un code permettant un autre niveau de contrôle de la transcription des gènes. Par exemple, l'acétylation de certains résidus permet l'interaction de différentes familles de coactivateurs (Fig 4) . À l'opposé, la méthylation à des endroits précis des queues des histones entraîne le condensement de la chromatine, le recrutement des histones H1 et parfois la méthylation de l'ADN, ce qui est un signe d'une région à faible activité transcriptionnelle (89,94). Ce «code des histones» est complexe et fait intervenir plusieurs modifications transcriptionnelles et plusieurs combinaisons de modification. Les éléments frontières discutés précédemment servant à délimiter des régions spécifiques de la

chromatine vont empêcher la propagation du signal transcriptionnel au-delà de la région ciblée par les récepteurs.

Les nucléosomes servent à la fois à réguler la transcription et à ordonner l'ADN dans une structure pouvant être compactée à un très haut niveau. La formation de la structure tridimensionnelle de l'ADN joue un rôle prépondérant et dynamique dans la transcription des gènes (99). Certaines données démontrent même que certains gènes peuvent sortir de leur territoire chromosomique pour être transcrit dans différentes régions «usines» du noyau (100,101).

3.2 Régulation de la transcription à longue distance

La complétion du projet de séquençage humain a permis de réexaminer le(s) rôle(s) de l'ADN ne codant pas pour des gènes classiques. Longtemps les histones ont été considérées comme statiques et possédant un rôle structural. L'étude des structures de la chromatine en trois dimensions rendue possible suite aux avancements technologiques, a permis de dégager qu'il existait des éléments de réponse aux facteurs de transcription situés loin des sites d'initiation de la transcription. Plusieurs théories ont vu le jour pour expliquer comment et pourquoi ces différents sites agissent sur de longues distances et l'implication potentielle de la chromatine.

Dans cette dernière section, les principales découvertes menant à la définition du concept tridimensionnel de boucles de chromatine et leurs effets sur la transcription seront abordées. Finalement, les observations menant à une réflexion sur le concept classique de promoteur, c'est-à-dire les quelques Kbp en amont du site d'initiation de la transcription, seront présentées.

3.2.1 La mise à jour du concept de promoteur

Les études chez la levure ont démontré que les sites de liaison des facteurs de transcription n'étaient pas limités aux promoteurs classiques, soit quelques milliers de paires de base. L'avancement d'études à grande échelle a permis de démontrer que le facteur de transcription Gcn4, possède près de 170 sites de liaison chez la levure, dont près de 120 sont situés à l'extérieur des régions promotrices (102). Cette association «lointaine» permet à Gcn4 de recruter la machinerie de modification de la chromatine sans recruter le médiateur (102), procurant une ouverture de la chromatine sur de longues distances augmentant ainsi l'accessibilité pour les autres facteurs régulant la transcription d'un gène donné. Ces observations démontrent que bien que le promoteur minimal soit requis, bien que la liaison de facteurs de transcription soit nécessaire pour apporter le médiateur et la machinerie transcriptionnelle de base, certains *enhancers* et leurs facteurs associés, peuvent se situer hors de la région «classique» et jouer un rôle de préparation à la transcription. Ceci rend inutilisable les techniques d'études classiques de promoteurs comme les vecteurs rapporteurs dû à la taille des vecteurs et à la difficulté d'introduire de tels vecteurs dans une cellule de mammifère.

Tout comme chez la levure, des études d'immunoprécipitation de la chromatine suivie d'une hybridation sur une puce d'ADN contenant différentes régions du génome (ChIP-CHIP) ont été réalisées avec différents facteurs de transcription de mammifères. Afin d'étudier les sites de fixation de manière exhaustive, les études ont été réalisées sur des puces d'ADN contenant la totalité des séquences non répétées des chromosomes humain 22. La première étude a été effectuée avec le facteur de transcription NF- κ B (103). NF- κ B peut lier des régions diversifiées sur le chromosome, en amont et en aval d'un promoteur, dans des régions codantes ou pas et même dans des introns, en conservant cependant un enrichissement marqué à proximité des sites d'initiations de la transcription des gènes. De façon étonnante, 40% des sites de liaison de NF- κ B se retrouvent à proximité de gènes n'étant pas régulés par la signalisation de TGF α (NF- κ B médie la signalisation de TGF α). Une étude suivante portant sur le facteur de transcription CREB démontre certaines

similarités conceptuelles avec NF- κ B (104). Les sites de liaison de CREB sont répartis sur la totalité du chromosome 22, autant dans les introns, des régions annotées que des régions non connues pour être transcrites. Contrairement à NF- κ B, les sites de fixation de CREB sont peu présents à proximité des sites d'initiations de la transcription. Ces observations ont remis en question la description classique de CREB comme facteur de transcription fixé à proximité de la boîte TATA et ont permis de réévaluer le «consensus» de fixation. Peu de cibles de la forskoline (composé qui médie une signalisation via CREB) possèdent un élément à proximité de l'initiation de la transcription du gène.

Ces dernières études à grande échelle, couplées à des analyses bioinformatiques confirment bien les observations réalisées chez la levure, à savoir que le concept minimaliste de promoteur est à réévaluer. Les facteurs de transcription peuvent lier partout sur le génome, peuvent être fixés à l'ADN à proximité ou pas des sites d'initiations de la transcription, et donc un gène cible d'un facteur de transcription ne possède pas nécessairement un site de fixation proximal. Dans le domaine des récepteurs nucléaires, des études portant sur le récepteur des androgènes ont permis de démontrer par immunoprécipitation des complexes liés à la chromatine qu'un élément situé à 4000 pb en amont du site d'initiation de la transcription serait un enhancer fonctionnel (105). Ceci laisse présager que les récepteurs nucléaires, qui ne sont qu'une classe particulière de facteurs de transcription, peuvent eux aussi agir sur la transcription via de grandes distances.

3.2.2 Les boucles de chromatines

Depuis plusieurs années, des études réalisées chez la drosophile et les levures ont démontré que la chromatine est dynamique et participe à la régulation des gènes (106). L'activation ou la répression des gènes, par des éléments lointains, en *cis* (même chromosome) et même en *trans* (chromosomes différents), a d'abord été caractérisée chez la drosophile (107,108). La formation de boucles de chromatine à l'intérieur d'un même chromosome a été observée de différentes façons chez la drosophile (108). Le gène de la protéine hsp70 est flanqué de deux éléments frontière *insulator* (scs et scs') liant chacune un facteur différent (ZNF ou BEAF32). Des études d'immunoprécipitation de la chromatine ont déterminé que les protéines qui sont normalement retrouvées chacune sur leur élément scs pouvaient se retrouver sur l'un ou l'autre des sites frontières lors d'un choc thermique (108-110). Une autre étude réalisée chez les embryons de drosophile a permis de détecter la présence d'une boucle de chromatine à l'aide de la technique de 3C (*chromosome conformation capture*) (Fig 7). Cette technique est basée sur le principe d'une réticulation par la formaldéhyde de complexes ADN-protéines, digestion de la chromatine et ligation en conditions de dilution limite. Ces techniques ont permis de démontrer que les deux régions «scs» se retrouvent à proximité (108-110). D'autres études sur le rétrotransposon *gypsy* réalisées par FISH ont démontré la présence d'une boucle de plus de 400 Kpb (108,111,112). Finalement des interactions entre les chromosomes (*trans*-) ont été démontrées pour le facteur de transcription GAGA, qui peut dans un même temps former des boucles en *cis*- et en *trans*- grâce à une oligomérisation (107,108). Ces quelques exemples démontrent l'existence des boucles de chromatines et ont permis d'élaborer des hypothèses face au(x) rôle(s) potentiel(s) de ces boucles, qui pourrait être d'isoler des éléments d'ADN de leur région immédiate et les amener à un autre endroit.

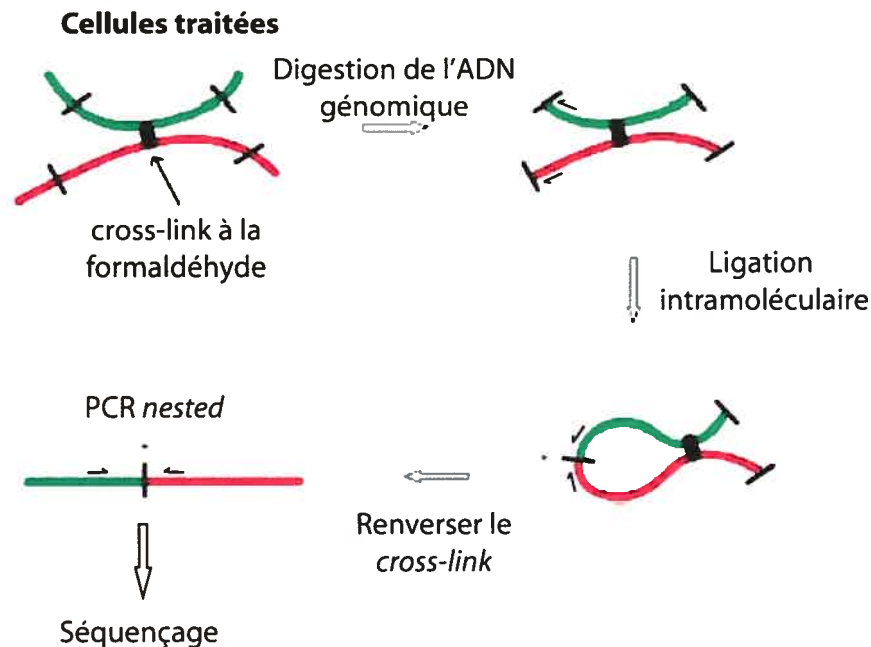


Figure 7 : La technique 3C : *Chromosome Conformation Capture*. La technique de 3C permet la détection d'une proximité entre différents endroits de la chromatine (ici par exemple deux régions, rouge et verte). En premier lieu, on lie de façon covalente la chromatine à un instant « t » suite à un traitement avec l'aide d'un agent de réticulation, ici la formaldéhyde. Par la suite, l'ADN génomique est extrait et digéré avec des enzymes de restriction. Plus l'enzyme coupe fréquemment dans le génome, plus les fragments récupérés par la suite seront petits. On dilue l'ADN digéré afin de permettre seulement des ligations intra-moléculaires. Après avoir renversé les liens covalents, on peut détecter les régions de ligation par PCR. On devrait retrouver par séquençage une partie des deux régions (illustrées ici par les couleurs rouges et vertes) qui se retrouvent à proximité dans notre extrait cellulaire suite au traitement. Figure adaptée de(113)

Ces observations ont donné lieu à différentes hypothèses sur l'utilité de ces éléments distants et leurs modes d'action dans la régulation de la transcription (93,106). Les éléments éloignés (enhancers et LCR) peuvent servir de lieu d'ouverture de la chromatine. Les nucléosomes peuvent être remodelés et modifiés par le recrutement de différents complexes protéiques via les facteurs de transcription. Il se produit alors une décondensation de la chromatine permettant aux protéines de la machinerie transcriptionnelle de se lier et de rechercher les sites d'initiation situés en aval de cet élément de régulation lointain. Le problème majeur de cette théorie est qu'elle ne peut pas expliquer la régulation en *trans* par un site de fixation (106). La deuxième théorie la plus probable est la formation de boucles de chromatine dans l'ADN qui permet aux sites de

liaisons lointains (*cis* ou *trans*) d'être relocalisés près d'une région où une concentration de corégulateur est nécessaire (ex: à proximité du site d'initiation de la transcription) ou bien pour les maintenir éloignés d'une région donnée (ex :site d'initiation d'un gène non transcrit) (106,114).

Cependant, il fallut attendre 2002 pour que deux groupes (115,116) démontrent *in vivo*, avec deux techniques très différentes, que les boucles de chromatines existent chez les mammifères et qu'elles ont clairement un rôle à jouer *in vivo*. La fonctionnalité des éléments LCR du gène de la β -globine de souris a été étudiée par le premier groupe. Ces différents sites se situent à plus de 30 Kpb des sites d'initiation de la transcription de la β -globine chez la souris et sont caractérisés comme étant hypersensibles aux coupures par la DNase (106,108,115,116). Puisque la transcription de la globine change durant la vie fœtale, la compréhension de la transcription de ce locus est très intéressante. Le premier groupe a adapté la technique de FISH pour permettre l'incorporation de biotine au site d'initiation de la transcription, de visualiser les sites de transcription et de les purifier pour quantification subséquente (RNA-TRAP) (115). Ils ont déterminé que le site hypersensible 2 (HS2) situé normalement à 50 Kpb se retrouvait à proximité du gène de β -globine fraîchement transcrit dans des embryons de souris (115). Le deuxième groupe a adapté une méthode utilisée chez la levure (113) pour détecter des interactions sur de longues distances. En utilisant la technique de 3C sur des embryons de souris, ils ont démontré que les éléments LCR du gène de la β -globine interagissent via des complexes protéiques au site d'initiation du gène (116). Ils ont de plus démontré clairement que ces boucles ne se forment pas dans les tissus où les gènes de la globine ne sont pas transcrit (116).

Par la suite, plusieurs études chez les mammifères ont permis de démontrer la présence de boucles de chromatine dynamiques, donnant ainsi un niveau de hiérarchie à la structure de la chromatine (106,108,117-119). Il a également été démontré chez un gène (locus cytokine TH2 et TH1) que la régulation en *trans* était possible chez les mammifères (100,117). La formation de boucles de chromatine dynamiques réconcilie les deux théories (boucle vs ouverture à longue distance); les éléments lointains, bien qu'ils se retrouvent à proximité du site d'initiation de la transcription, recrutent toujours des complexes de protéines servant à ouvrir la chromatine. Ceci est en accord avec le modèle où les LCR peuvent réguler la dynamique de l'ouverture de la chromatine (93,106,108,114,117,120).

La distribution de site de fixation des facteurs de transcription dans l'ensemble du génome et la formation de boucles de chromatine participant à la transcription mettent en lumière la nécessité d'élargir le concept de promoteur tant en amont qu'en aval du site d'initiation de la transcription. L'approfondissement des connaissances à l'échelle génomique des sites de liaison des récepteurs nucléaires, la caractérisation des gènes régulés par les oestrogènes dans différents types cellulaires et la compréhension des mécanismes de formation des boucles de chromatine de façon tissus-spécifiques apporteront des précisions inestimables sur la signalisation hormonale.

Chapitre 4 : Problématiques

Au Canada, une femme sur 9 développera un cancer du sein au cours de sa vie. Ce dernier est la forme de cancer la plus répandue chez les femmes. Malgré tout, l'avènement de méthodes de détection efficaces et l'amélioration des traitements ont permis de diminuer de façon importante le taux de mortalité depuis le début des années 90. (www.cancer.ca)

Les oestrogènes jouent un rôle très important dans le développement des tumeurs mammaires. Environ les deux tiers des tumeurs primaires sont positives pour l'expression du récepteur ER α (121). La compréhension du rôle et des modes d'action du récepteur est donc un enjeu très important dans la recherche sur le cancer du sein. Des anti-oestrogènes totaux et partiels sont utilisés en clinique depuis déjà plusieurs années pour traiter le cancer du sein. De plus, des inhibiteurs de l'aromatase peuvent aussi être utilisés en clinique de façon à inhiber la signalisation des oestrogènes. Ces médicaments, bien qu'aidant la régression des tumeurs et la prévention des rechutes, peuvent causer des effets secondaires importants dû aux rôles diversifiés des oestrogènes chez la femme et sont sujets à des mécanismes d'échappement.

Bien que plusieurs études se soient penchées sur la façon dont le récepteur lie l'ADN, il y a encore peu de gènes cibles directs identifiés comme ayant un rôle prépondérant dans la prolifération tumorale et très peu de sites de liaison directe de ER α sur l'ADN sont connus *in vivo*. Le récepteur peut réguler l'expression des gènes par liaison à l'ADN (étude exhaustive réalisée sur le gène TFF1 (53)), mais peut aussi passer via des mécanismes alternatifs en jouant un rôle de coactivateur (122,123) ou bien en étant impliqué dans les voies non génomiques (124-126). Quels mécanisme d'action de ER contribuent à la prolifération tumorale? Les mécanismes alternatifs jouent-ils un rôle prépondérant? Ou bien simplement, les méthodes utilisées jusqu'à présent pour la recherche

d'éléments de réponse sont-elles à trop faible échelle, ce qui explique le peu de sites de fixation caractérisés?

Afin de mieux comprendre les mécanismes tumoraux propres aux oestrogènes médiés par les ER, il est nécessaire d'identifier les gènes cibles primaires des oestrogènes. Dans un premier temps, l'approfondissement de la spécificité de la liaison à l'ADN des ER (Annexe 1) et la définition des sites de liaison de haute affinité (Chapitre 5) représentent des étapes préalables nécessaires pour faciliter les recherches des sites de fixation des récepteurs. Dans un deuxième temps, afin d'avoir une vue à grande échelle des sites de liaison de haute affinité des récepteurs des oestrogènes des recherches bioinformatiques seront réalisées sur l'ensemble du génome (Chapitre 5,6). Dans un troisième temps, la corrélation entre la présence d'une régulation oestrogénique pour un gène donné et la présence d'un site de fixation de haute affinité, permettra d'identifier les mécanismes de régulation des gènes cibles directs (Chapitre 6) dans le but de mieux comprendre les mécanismes de prolifération cellulaire. Finalement, l'identification dans nos recherches bioinformatiques de sites de fixation distaux du récepteur a nécessité l'approfondissement du mécanisme de régulation de la transcription par les récepteurs des oestrogènes (Chapitre 7).

Deuxième partie

Chapitre 5: *Genome-wide Identification of High Affinity Estrogen Response Elements in Human and Mouse*

Chapitre 6: *Mechanismes of Gene Regulation by Estrogen Receptor Alpha in Breast Cancer Cells.*

Chapitre 7: *Regulation of GREB1 Transcription by Estrogen Receptor Alpha Through a Multipartite Enhancer Spread Over 20 KB of Upstream Flanking Sequences*

Chapitre 5 : Genome-wide Identification of High Affinity Estrogen Response Elements in Human and Mouse

Véronique Bourdeau, **Julie Deschênes**, Raphael Métivier, Yoshihiko Nagai, Denis Nguyen, Nancy Bretschneider, Frank Gannon, John H. White, and Sylvie Mader.

Publié dans : *Molecular Endocrinology*, 2004 Jun;18(6):1411-27

En tant que deuxième auteure, j'ai effectué les études EMSA sur les différents ERE naturels identifiés dans la recherche. J'ai également assisté Dr Bourdeau à la réalisation des différentes tables de l'articles. J'ai finalement participé à la correction du manuscrit et des différentes figures.

Genome-wide Identification of High Affinity Estrogen Response Elements in Human and Mouse

Abbreviated title: Genomic search for near-consensus EREs

Département de Biochimie (V.B., J.D., D.N., S.M.), Université de Montréal, C.P. 6128 Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada ; Departments of Physiology (V.B., J.H.W.), and Medicine (J.H.W., S.M.), McGill University, 3655 Drummond St, Montreal, Québec H3G 1Y6, Canada; McGill University and Genome Québec Innovation Centre (Y.N.), 740 Docteur Penfield Avenue, Montreal, Québec H3A 1A4, Canada; European Molecular Biology Laboratories (EMBL) (R.M.*, N.B., F.G.), Meyerhofstrasse, 1. D-69117, Heidelberg, Germany.

Address all correspondence and requests for reprints to : Sylvie Mader, Département de Biochimie, Université de Montréal, CP 6128 Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada. E.mail: sylvie.mader@umontreal.ca.

*: Current address: Equipe EMR; UMR 6026 CNRS. Campus de Beaulieu, 35046 Rennes Cedex, France.

Key words: Estrogen receptor, estrogen response element, estrogen target gene, chromatin immunoprecipitation, genome-wide search.

Abbreviations: ADORA1, adenosine A1 receptor; C3, complement component 3; CASP7, caspase 7, apoptosis-related cysteine protease; CDS, coding sequence; ChIP; chromatin immunoprecipitation; COX7A2L, cytochrome c oxidase subunit VIIa polypeptide 2 like; CTSD, cathepsin D; EBAG9, estrogen receptor binding site associated, antigen 9; EFP, estrogen-responsive finger protein; ER, estrogen receptor; ERE, estrogen response element; FGF9, fibroblast growth factor 9; GAD2, glutamate decarboxylase 2; GAPD, glyceraldehyde-3-phosphate dehydrogenase; GREB1, gene regulated by estrogen in breast cancer protein; IGFBP4, insulin-like growth factor binding protein 4; KHDRBS3 (T-STAR/ETOILE) KH domain containing, RNA binding, signal transduction associated 3; LY6E, lymphocyte antigen 6 complex, locus E; NRIP1 (RIP140), nuclear receptor interacting protein 1; OVGP1, oviductal glycoprotein 1; P-PolIII, phosphorylated RNA polymerase II; RBM, RNA-binding motif; RNF14, ring finger protein 14; SCNN1A, sodium channel, nonvoltage-gated 1 alpha; TBP, TATA box binding protein; TERT, telomerase reverse transcriptase; TESK1, testis-specific kinase 1; TFF1/pS2, trefoil factor 1; TSHB, thyroid stimulating hormone beta; VitA2, *Xenopus* vitellogenin A2; ZNF147, zinc finger protein 147.

ABSTRACT

Although estrogen receptors (ERs) recognize 15 bp palindromic estrogen response elements (EREs) with maximal affinity *in vitro*, few near-consensus sequences have been characterized in estrogen target genes. Here we report the design of a genome-wide screen for high-affinity EREs and the identification of ~70,000 motifs in the human and mouse genomes. EREs are enriched in regions proximal to the transcriptional start sites, and approximately 1% of elements appear conserved in the flanking regions (-10 kb to +5 kb) of orthologous human and mouse genes. Conserved and non-conserved elements were also found, often in multiple occurrences, in over 230 estrogen-stimulated human genes previously identified from expression studies. In genes containing known EREs, we also identified additional distal elements, sometimes with higher *in vitro* binding affinity and/or better conservation between the species considered. Chromatin immunoprecipitation experiments in breast cancer cell lines indicate that most novel elements present in responsive genes bind ER α *in vivo*, including some EREs located up to ~10 kb from transcriptional start sites. Our results demonstrate that near-consensus EREs occur frequently in both genomes and that while chromatin structure likely modulates access to binding sites, far upstream elements can be evolutionarily conserved and bind ERs *in vivo*.

INTRODUCTION

Estrogens, such as 17- β -estradiol (E2), are steroid hormones produced mainly by the ovary and act by endocrine, paracrine and possibly autocrine mechanisms on a number of target tissues. Their actions on the reproductive, cardiovascular and central nervous systems, and on bone [4-7] are mediated by two estrogen receptors, ER α and ER β [7, 8], members of the nuclear receptor superfamily of ligand-inducible transcription factors [9-12]. A central DNA binding domain, which corresponds to the conserved C region [13, 14], is responsible for primary sequence recognition and cooperative binding by receptor dimers on elements containing two appropriately spaced copies of specific recognition motifs. The ligand binding domain contains a strong dimerization interface that stabilizes receptor homodimers or heterodimers [15-17] and enhances binding to imperfect motifs and even to single copies of recognition motifs [14, 18-20]. However, ERs bind with highest affinity to 15 bp palindromes composed of PuGGTCA motifs separated by 3 variable bp [20-22]. These estrogen response elements (EREs) are recognized by ERs with high specificity, since other nuclear receptors bind either different motifs or similar motifs with different spacing and/or orientation [10, 22-24]. Studies using chromatin immunoprecipitation approaches have demonstrated that ER binding to EREs is hormone-inducible *in vivo* [25-28] and results in the ordered recruitment of a series of coactivator complexes, leading to histone acetylation, chromatin remodeling, and enhanced recruitment of the basal transcription machinery [29-36].

Surprisingly, only a few near-consensus EREs have been characterized to date in the promoters of E2-regulated human or mouse genes. This probably reflects in part the fact that ERs can mediate estrogenic regulation through multiple mechanisms, including tethering to DNA via protein-protein interactions with other transcription factors [22, 37, 38], or so-called "non-genomic" effects, which affect gene expression by modulating the activity of upstream components of various intracellular signaling pathways [39-41]. The paucity of known near-consensus EREs may also result from the fact that their presence is

usually investigated only in relatively proximal sequences of E2-responsive genes. However, with sequences of the human and mouse genomes now available, genome-wide screening strategies can be used to expand the repertoire of potential high-affinity ER-binding sites.

In this study, we have sought to identify high affinity EREs in the human and mouse genomes, and to determine whether these elements represent *bona fide* ER binding sites *in vivo*. Indeed, little is known about the effect of chromatin structure or promoter context on ER binding *in vivo*. In addition, it is not clear how far from transcription start sites EREs can be positioned to recruit ERs and play a role in regulating gene expression. Our characterization of high-affinity EREs in the human and mouse genomes will greatly facilitate the study of the mechanisms modulating accessibility of ERs to their binding sites and of the subsequent events leading to transcriptional regulation by estrogen. This analysis will also provide a powerful resource for researchers analyzing the molecular events underlying the broad-ranging physiological actions of estrogens.

RESULTS

Identification of sequences corresponding to high-affinity ER binding sites *in vitro*.

Palindromic response elements composed of two PuGGTCA motifs separated by 3 bp, such as the one found in the *Xenopus* vitellogenin A2 (VitA2) promoter, represent the highest affinity binding sites for estrogen receptors *in vitro* (Fig. 1A) [20-22]. Surprisingly, only three perfect palindromic EREs have been identified to date in the vicinity of E2-regulated genes in the human genome. These genes, EBAG9, COX7A2L and EFP/ZNF147, were cloned by screening of a CpG island genomic library for binding to estrogen receptors [42-44]. In addition, a few near-consensus EREs have been characterized in human E2-responsive genes (see Table 1 for EREs diverging from consensus at one or two positions). A systematic analysis of single nucleotide substitutions in gel shift assays indicates that all elements with one variation still bind estrogen receptors *in vitro* (Fig. 1B). However, some

replacements were more detrimental than others (-5C, -4C, -2A), reflecting steric or charge clashes between specific ER amino acids and bases of the variant EREs (Nguyen et al., in preparation). Note that, as expected from the high degree of homology between the DNA binding domains of ER α and ER β (90% in region C as defined in [13]), all mutations had a similar impact on complex formation by the two receptors. Introduction of symmetrical replacements in both arms of the palindrome generally affected ER binding much more drastically than when a unilateral change was introduced, although a few replacements were well tolerated even when present in both arms (Fig. 1C). Interestingly, examination of the known natural elements containing two replacements indicates that they all contain one change with mild effect on affinity (-6Py or -1G; Fig. 1B, C and Table 1) combined with another replacement.

Screening the human and mouse genomes for high-affinity EREs.

Based on the above-described observations, we decided to identify all occurrences of EREs conforming to consensus, or with one variation, or with two variations including either -6Py or -1G. This defines 588 different sets of 15 bp sequences with random 3 bp spacers. 71,119 elements were identified in the human genome and 65,012 in mouse. Note that the elements containing CpG dinucleotides were found to be largely underrepresented in both genomes (data not shown), consistent with the lower genomic frequency of this dinucleotide due to deamination of the methylated C [45]. In particular, symmetrical palindromes containing PuGGTCG motifs (2 CpGs), which represent good ER-binding sites (Fig. 1C), were found only 14 times in the human genome, whereas consensus palindromes were represented 891 times. The distribution of consensus EREs among the chromosomes is generally consistent with chromosomal size in both genomes (Fig. 2A, B). While perfect EREs were not clustered together, several imperfect palindromes were found to be part of tandem repeat elements at different chromosomal locations in the human and mouse genomes, increasing the frequency of these elements (Bourdeau V. et al., in preparation).

The numbers of near-consensus EREs found in the human and mouse genomes suggest that functional EREs should be found in the vicinity of genes more frequently than reported to date. Note that most previously identified EREs are located within 0.5 kb of transcriptional start sites (see Previously characterized elements in Table 1, with the exception of the TERT and the EFP/ZNF147 EREs). It is not clear whether this bias reflects a functional requirement for proximal EREs or the fact that searches for EREs in E2 target genes usually focus on promoter sequences immediately upstream of the transcriptional start site. Using a cut-off distance of -10 to +5 kb from transcriptional start sites, we found EREs in the vicinity of 17,353 transcriptional start sites in human, and 15,284 in mouse (see website at URL: <http://mapageweb.umontreal.ca/maders/eredatabase/>) Note that the number of genes containing EREs is slightly smaller (12,515 in human and 11,810 in mouse), as there can be several mRNAs and more than one ERE per gene. We further examined the distribution of EREs in the -10 to +5 kb region with respect to the 5' end of mRNAs (whenever several mRNAs are mapped, distances were calculated to the most 5' initiation site). Note that EREs found close to more than one gene were counted several times for characterization of the distribution with respect to transcriptional start site. Surprisingly, EREs were more abundant in the 0 to +1 kb region around transcriptional start sites in both genomes (excess of about 25%), while the number of elements found upstream of -1 kb or downstream of +2 kb were usually near or below the average number of near-consensus EREs found per kb between -10 and +5 kb (Fig. 2C). Compared to the frequency expected from random distribution of all EREs found in the genome, the number of individual EREs in the -1 to +2 kb region of 5' mRNAs was higher than expected by 64% in human and 29% in mouse (data not shown).

These data suggest that the distribution of near-consensus EREs reflects mostly a random distribution of these sequences in the different chromosomes, with enrichment in the close vicinity of transcriptional start sites. The large total number of EREs located in the vicinity (-10 to +5 kb) of transcriptional start sites may suggest that these elements can mediate regulation of a much larger fraction of the genome in different estrogen target

tissues than previously recognized. However, some of these elements may not represent binding sites *in vivo*, possibly because of chromatin accessibility, and/or may not participate in transcriptional regulation.

Identification of estrogen response elements conserved in human and mouse.

Near-consensus EREs previously identified in human genes often have counterparts at similar positions in their mouse orthologs (Table 1). Because conservation of newly identified EREs between the two species may indicate a functional role of these elements, we searched for the presence of EREs located in known human and mouse orthologs with less than 2 kb difference in distance from their respective transcriptional start sites (see Materials and Methods). EREs in 660 different pairs of orthologs were thus identified (Supplementary online Table 1). As expected from the distribution of EREs in the vicinity of genes, conserved EREs were also more abundant in the 0 to +1 kb region after the transcription start site (Fig. 2C; note that this analysis was intentionally limited to the -8 to +3 kb region. The numbers of conserved EREs found between -10 and -8 kb or +3 and +5 kb would appear artificially low, since a conserved ERE found at a distance of 2 kb or less in the other species may not be included in the -10 to +5 kb window from the transcriptional start site). However, conserved EREs were proportionately more represented in the 0 to +2 kb region than bulk EREs, while elements in the -5 to -8 kb region were markedly less represented (Fig. 2C). Nevertheless, we note that a significant number of conserved EREs mapped between -5 and -10 kb of the transcriptional start sites (24.6%). It remains possible that some of these distal EREs may be located closer to another gene, or even within an open reading frame, accounting for their conservation.

To investigate whether location of EREs in coding sequences (CDS) may account for their conservation, we characterized the position of these elements for a subset of conserved EREs, i.e. those that were identical in sequence between the two orthologs (Table 2). A total of 41 elements were found in the vicinity of 47 genes. Out of these elements, 14 were located fully and 4 partly in a CDS (Table 2 bottom and middle,

respectively), while 23 were not found in any annotated ORFs (Table 2, top). More than half of these (12 EREs) were located between -1 to $+1$, another 7 between -1 and -10 kb, and another 4 between $+1$ and $+5$ kb. Thus these results indicate that EREs that are conserved independently of their presence in a CDS are more likely to be found close to the 5' end of mRNAs, although some elements are still located at several kb of the start sites. Further, the strong conservation of these elements, which is unlikely to result from chance alone, may indicate a role in the regulation of the neighboring genes.

Three known E2 target genes were present in the set of genes containing perfectly conserved EREs. The EBAG9 and COX7A2L genes contain previously identified EREs (Table 1). The FGF9 gene is also a known E2 target [46], although the presence of a near-consensus ERE in the proximal promoter was not previously reported. FGF9 mediates the proliferative action of estrogen on stromal endometrial cells [46]. It also plays a role in glial cell development [47] and testicular embryogenesis [48]. Its functional inactivation in mouse leads to male-to-female sex-reversal [48]. Interestingly, several other genes containing fully conserved EREs in their flanking sequences (Table 2) are linked to testicular function. TESK1 is expressed in testicular germ cells in a developmentally-regulated manner and is thought to play a role in spermatogenesis around the stages of meiosis and/or early spermiogenesis [49, 50]. KHDRBS3 (T-STAR/ETOILE) is a SAM68 homolog that is expressed primarily in the testis and interacts with RBM, an RNA-binding protein implicated in spermatogenesis [51]. FGF8 is an androgen-induced gene expressed in testes specifically in prespermatogonia during fetal days 16 to 17 [52]. Finally, RNF14 (ARA54) is an androgen receptor coactivator expressed in testis and prostate [53]. Although estrogenic regulation of these genes remains to be demonstrated, these observations are noteworthy in view of the fact that estrogen deficiency in aromatase knock-out models or genetic ablation of the ER α gene leads to defects in spermatogenesis and male sterility [54-56].

Proximal and distal EREs in known estrogen responsive genes.

To assess the potential of EREs found in this screen to bind ERs, we first compared our list of elements lying between -10 and +5 kb of the 5' ends of known genes to published data on E2-upregulated genes identified through expression studies in different human target tissues. This generated a list of 345 elements close to 236 genes (Table 3). Several of these genes were identified from a gene array experiment performed in the presence of cycloheximide, suggesting direct regulation [57]. Interestingly, several promoters already shown to contain functional EREs were found to include additional elements, sometimes located more distally (Table 3, see also selected elements in Table 1). When assayed for binding to ER α in gel shift assays, several of these novel elements bound similarly or even better than those previously characterized (Fig. 3A, OVGP1, CTSD and TERT). Of interest, the mouse ortholog of the CTSD gene did not contain a near-consensus element corresponding to the known human proximal element, but our search identified a new element located 8 kb upstream that is conserved between human and mouse. Another example of gene containing conserved distal EREs is the EFP/ZNF147 gene. Our screen (performed in this case with a larger cut-off) identified a conserved imperfect element in the 3' UTR of the mouse and the human orthologs, different from the consensus ERE previously reported in the human gene (Table 1). The presence of distal elements in E2-regulated genes suggests that they may participate in the estrogenic regulation of these genes.

Some estrogen-inducible genes for which EREs had not been characterized were found to contain several high-affinity elements. For instance, the GREB1 (gene regulated by estrogen in breast cancer) gene contains three consensus EREs located at -1.5 kb, -9.5 kb and -21 kb (Table 1). In addition, the human CASP7 gene was found to contain 7 tandem repeats of a near-consensus ERE within the proximal 500 bp (Table 1). The presence of multiple elements raises the possibility of synergism at the level of either DNA binding or transcriptional activation in the regulation of these genes by estrogen, and may suggest

propagation of chromatin changes induced by recruitment of ER coactivators over a large distance.

Functional analysis of newly identified EREs.

To assess whether EREs selected among those found in the promoters of E2-responsive genes (Table 1, bottom) bind ER α *in vitro* with the predicted high affinity, gel shift assays were performed with elements in their natural context (15 bp core with 6 bp of flanking sequences; Fig. 3B). Although flanking sequences slightly modulated binding compared to the VitA2-derived elements carrying the same core ERE sequence, all elements were found to bind ER α efficiently. Chromatin immunoprecipitation (ChIP) experiments were performed next to monitor *in vivo* recruitment of ER on the newly identified EREs. These experiments were conducted in ER α -positive MCF7 cells, ER α -negative MDA-MB231 cells or MDA-MB231 cells stably transfected with ER α (MDA::hER α). As positive controls for ER α recruitment, both the pS2/TFF1 and C3 elements (Table 1) were used and, as expected, demonstrated E2-dependent recruitment of ER α on their proximal promoters (Fig.4). We investigated binding of ER α to 14 EREs found in 10 genes known to be regulated by estrogen in breast cancer cells (ADORA1, CASP7, CTSD, GAPD, GREB1, IGFBP4, LY6E, NRIP1, SCNN1A, TERT; Table 1 and refs within). In the two ER α -expressing cell lines, E2-induced ER α binding was observed on 13 out of 14 of the elements (which include both conserved and non conserved EREs, see Table 1), except for the GAPD ERE where binding was restricted to MCF7 cells (Fig. 4). No binding was observed in the parental MDA-MB231 cells, or with pre-immune rabbit IgG, on any of the promoters tested. Consistent with the published regulation of these genes by estradiol in breast cancer cells, association of the fragments of genomic DNA around the transcriptional start sites with the TATA box binding protein (TBP) and the phosphorylated polymerase II (P-PolII) was induced in the presence of E2 (Fig. 4). Notably, in addition to the proximal elements (ERE2) found in the CTSD and TERT genes, both upstream elements (ERE1) bound ER α . Binding of both the GREB1 EREs at -9.5 kb (ERE2) and at -1.5 kb (ERE3) was also observed, but the most upstream ERE (-21 kb, ERE1) did not bind

in spite of its consensus ERE sequence. This may suggest that this element was not accessible for ER binding, possibly because of chromatin organization.

Binding of ERs to EREs located in two genes that are not known to be estrogen targets in breast cancer cells was also investigated (GAD2, TSHB, Fig. 4). In contrast to results obtained with other ERE-containing promoters, ER α binding was not detected on the GAD2 ERE. Moreover, no binding of TBP or P-PolIII was observed in the absence or presence of hormone. GAD2 (glutamic acid decarboxylase 2) catalyses the synthesis of the inhibitory neurotransmitter γ -amino butyric acid, and its expression is restricted to neural tissues and pancreatic islets [58]. Its regulation by estrogen has been described in rat and goldfish brain [59, 60]. However, no expression of GAD2 was detected in breast cancer cells by RT-PCR analysis (data not shown). Absence of ER binding to these sites in breast cells may reflect limited access due to chromatin conformation. On the other hand, ER α bound to the TSHB (thyroid stimulating hormone beta) ERE whereas no binding of TBP and P-PolIII was detected either on the reported start site [61] or on another site described in NCBI AceView (Fig. 4). Note that regulation of TSHB by estradiol has been observed in mouse pituitary [62]. It is possible that another transcriptional start site may be used in MCF7 cells, or that this gene is inactive in these cells in spite of ER α binding to upstream sequences.

Results from ChIP experiments demonstrate that it is possible to identify *bona fide* ER binding sites both in proximal and distal promoter sequences of E2 target genes, suggesting that the search for ER binding sites should not be limited to promoter sequences immediately upstream of the transcriptional start site. In addition, our results indicate that access to some high affinity binding sites is restricted *in vivo*, possibly in a tissue-specific manner.

DISCUSSION

In this study, our goal has been to define high affinity *in vitro* ER binding sequences and to characterize the occurrence of these motifs throughout the human and mouse genomes. This approach differs from those of previous studies, which have aimed at characterizing a larger array of potential binding sites for ERs or nuclear receptors in defined stretches of genomic sequences [63]. These approaches are based on nucleotide frequency matrices constructed via compilation of the relatively small number of known natural response elements, and can identify many variant sequences. Although the sensitivity of the detection can be lowered for screening of genomic sequences, these types of detection programs cannot be easily used for genome-wide studies (the maximal length of input sequences is about 100 kb for Dragon ERE Finder v.2.0 and 30 kb for the NUBIScan program). On the other hand, we chose to identify only high affinity binding sites in a genome-wide approach in order to determine their distribution with respect to transcriptional start sites, their conservation between the mouse and human genomes, and their functionality as receptor binding sites *in vivo*. Our choice of elements was based both on functional validation of variant EREs by *in vitro* binding assays, and on the variations observed in known near-consensus sequences. While not within the scope of this study, our functional characterization of EREs both *in vitro* and *in vivo* is also expected to contribute to the development of refined nucleotide frequency matrices for the detection of a wider range of elements in genomic sequences of interest.

Our screen identified a large number of EREs (71,119 in the human genome, or about one every 43 kb of genomic DNA). Not surprisingly, the frequency of ERE occurrence was found to be highly dependent on their sequence. For instance, elements containing CG dinucleotides are drastically less represented than other EREs. CG dinucleotides are represented at a frequency of about 0.8% in the genome, which is 5 times less than the expected frequency based on the typical fraction of Cs and Gs [45]. Elements containing two CG dinucleotides are found at only about 1-2% of the number of consensus

EREs, which do not contain CG dinucleotides except possibly in the spacer. On the other hand, consensus EREs are slightly more represented than the expected frequency for representation in random sequences (122% in human and 150% in mouse).

The sequence bias for some of the elements and the observation that some high affinity EREs are found in tandem repeats precluded evaluation of which proportion of near-consensus EREs are under positive selective pressure, as could be expected if they mediate important regulatory roles. However, it was possible to evaluate the relative distribution of EREs in the vicinity of human or mouse genes. We have calculated that the number of EREs found in the human genome in the -1 to $+2$ kb region around the 5' end of mRNAs exceeds the number of elements expected from their frequency in the genome by 64%. It is unclear whether this represents conservation of elements found in the vicinity of transcriptional start sites, or simply bias for undetermined causes. A contributing factor may be the increased GC content in CpG islands, which are associated with the 5' end of genes. Consensus EREs have an average GC content varying between 40-73% (depending on the identity of the variable bases). For most sequences, this is higher than the genome-wide average of 41%, but compatible with the 60-70% average in CpG islands [45]. It is however surprising that the main peak of ERE representation is in the 0 to $+1$ kb region around the 5' end of mRNAs both in human and mouse. While the basis for this observation is unclear, it is worth noting that only the most 5' mRNA start site for each gene was taken into consideration in this representation. Thus, some of these EREs may actually belong to promoters of downstream initiation sites.

We performed next a direct comparison of elements found in human and mouse orthologs to assess the fraction of total EREs that is conserved in the two species. 660 out of the 9944 known orthologs (Mouse Genome Informatics database) contained one or several conserved EREs at similar relative positions in the 5' flanking sequences, i.e. at 2kb from each other (708 conserved elements in total). This corresponds to about 1% of total elements in the human or mouse genome. Note that our criteria for conservation are

relatively stringent, including both limited sequence variations (no more than 2 differences from consensus) and positional constraints (relative positions of EREs not further apart than 2 kb). Other possible sources of under-representation of conserved EREs are the incomplete identification of human/mouse orthologs, and the lack of systematic annotation of the mRNAs 5' ends in the human and mouse genomes. Nevertheless, we have estimated that there is an ~74% over-representation of EREs compared to chance occurrence of these elements in both the human and mouse orthologs within 2 kb of each other (708 versus 407 elements, see Materials and Methods). This over-representation suggests functional conservation of EREs throughout evolution, although we cannot rule out that functions unrelated to recruitment of ERs may contribute to this conservation (see below). The distribution of the conserved EREs was similar to that of total EREs in the vicinity of human or mouse genes, but displayed a more marked concentration in the vicinity of the transcriptional start sites (see Fig. 2C). Interestingly, the percentage of conserved elements dropped markedly upstream of -5 kb, suggesting that upstream EREs are less likely to be conserved than elements closer to the initiation site, albeit these EREs still represent 24% of the total number of conserved elements between -10 and +5.

Apart from conservation due to functional importance, the most likely reason for preservation of an ERE is its location within the coding sequence of a gene. This parameter might contribute to the over-representation of EREs located downstream of the initiation start sites. In addition, some upstream elements may also be found in the coding sequences of other genes. The possible contribution of coding sequences in preserving ERE motifs in the human and mouse genomes was examined in the extreme case of EREs found to be totally conserved in sequence between the two species, including the Pu/Py and spacer bp. Of 41 distinct EREs flanking 47 genes, 14 were found in coding sequences in both species and 4 additional elements overlapped a CDS. It is unclear whether these elements could function as binding sites *in vivo*. Binding of yeast transcription factors such as Gcn4, Sbf, Mbf and Rap1 to elements within open reading frames occurs reportedly less frequently than in promoter sequences [64], and in the case of Gcn4 results in recruitment of HAT and

SWI-SNF coactivator complexes, but not of the Mediator complex [65]. Finally, 23 EREs were not found in coding sequences in either species. Interestingly, more than half of these EREs (12 elements) were located between -1 to +1 kb of the transcriptional start site, including 8 elements between 0 and +1 kb. Therefore, while the presence of EREs in coding sequences may account in part for element conservation, it does not explain the large proportion of elements found immediately after the annotated transcriptional start site. Of interest in this regard, a recent publication indicated that an important fraction (40%) of NF κ B *in vivo* binding sites in chromosome 22 is located in intronic sequences [66].

Even though EREs appear less represented and/or conserved when located farther upstream of the transcriptional start site (-1 to -10 kb), our results suggest that they may contribute to estrogen target gene regulation. Indeed, some of these distal EREs are conserved between the human and mouse genomes and bind ER α *in vivo*, as demonstrated by our ChIP experiments. These observations validate the choice of the -10 to +5 kb window around transcriptional start sites used in this study to identify potential ER binding sites. However, the frequency of near-consensus EREs found within this range strongly suggests that not all of these elements mediate transcriptional regulation of neighboring genes in a given cell context. It is unclear at this point whether access to some of these elements is restricted, or whether bound receptors may be unable to transactivate neighboring genes. The number of EREs in bulk genomic DNA also indicates that if all elements were accessible, binding sites would likely outnumber the molecules of receptor dimers in estrogen target cells. Indeed, the number of estrogen binding proteins in rat uterine cells was estimated via either biochemical or autoradiographical methods to vary between 5000-30000 molecules per cell [67-70]. Large-scale studies of transcription factor binding sites using ChIP assays have also indicated that the number of *in vivo* binding sites for different transcription factors in the genome is high. For instance, in ChIP experiments performed in HeLa S3 cells, 15% of the genes in chromosome 22 were found to contain functional NF κ B binding sites within 10 kb [66]. In addition, 11% of the promoters were found to contain one or several high affinity c-myc binding sites, suggesting competition

between target sites in chromatin for limiting amounts of Myc protein levels [71]. Alternatively, another possibility that arises from these observations is that stochastic and/or dynamic binding of transcription factors may allow usage of a larger number of binding sites.

We are aware that the total number of potential ER binding sites is even much higher than reported here. Elements not included in our search may represent good binding sites *in vivo*, including those that display a relatively low affinity *in vitro*. While the compiled list of EREs is clearly not exhaustive, our ChIP experiments support the notion that the near-consensus EREs identified in estrogen-target genes represent likely binding sites *in vivo* (13 out of 14 elements bound). Nevertheless, access to some of these elements may be restricted irrespective of their strength as ER binding sites *in vitro*. Neither a perfect ERE located within about 20 kb of the transcriptionally active GREB1 estrogen target gene, nor a near-consensus ERE located in the non-transcribed GAD2 gene were bound in ChIP experiments. Although results obtained in two breast cancer cell lines were very similar, it will be of interest in future studies to examine further how binding site recognition and coactivator recruitment is affected by cellular context. The future availability of large-scale gene expression studies performed in different tissues will allow to expand the list of known E2 target genes and to determine whether the near-consensus EREs in their flanking sequences are bound in a tissue-specific manner. In addition, our data will facilitate future studies comparing the patterns of coactivators recruited by distal/proximal elements. For instance, it is possible that upstream EREs would recruit HAT and/or SWI/SNF complexes, resulting in long-range opening of chromatin and facilitating access of enhancer proteins to far upstream flanking sequences, while only proximal promoter sequences may be able to recruit Mediator complexes. Alternatively, chromatin loops may allow upstream elements to participate to the recruitment of the basal machinery on the transcriptional start site. It is our hope that the database of high affinity EREs derived from this study ([URL: http://mapageweb.umontreal.ca/maders/eredatabase/](http://mapageweb.umontreal.ca/maders/eredatabase/)) will prove a useful tool for the characterization of primary E2-regulated genes in various human and mouse target tissues,

and will ultimately enhance our understanding of the molecular mechanisms underlying the physiological actions of estrogens.

MATERIALS AND METHODS

Bioinformatics

The algorithms developed (Y. Nagai et al., in preparation) enable a search of the NCBI fasta and gbs files of the Human Genome (Built 33 - June 6, 2003) or of the Mouse Genome (Built 30 - March 11, 2003) for a specified group of sequences, and extract the positions of matching motifs in the genome contigs as well as the coordinates of the surrounding genes, mRNAs and CDS within a preset cutoff distance of each motif. The programs, written in C, were run on an SCI Origin 2000 with 32 CPUs IP 27, R10000, 300 MHz processor (16 Go) using UNIX IRIX 6.5. Results presented in this article were generated using a cutoff of -10 to + 5 kb of the mRNA 5' ends (database available at <http://mapageweb.umontreal.ca/maders/eredatabase/>). Statistics for the total number of EREs occurring in the genomes were derived from lists of elements generated before the cutoff was applied. The total number of consensus EREs in the human genome was 891, and the number in the mouse genome was 923. Expected frequency in random DNA sequences was calculated as the total number of bp in the genome divided by the frequency of occurrence of a sequence with specified bp at 10 positions and two bp choices at two positions ($3,069,334,246/4^{11} = 732$ EREs in the human genome, $2,578,250,392/4^{11} = 615$ EREs in the mouse genome).

Conserved EREs were identified as elements present in both human and mouse gene orthologs (listed at the Mouse Genome Informatics database) at distances comprised between -10 to + 5 kb from their respective mRNAs 5' ends, and differing by less than +/- 2kb. As in numerous instances transcriptional start sites are mapped at the ATG codon in at least one species, or at the 5' end of the gene for alternative upstream start sites, we have included EREs present in orthologs which were distant from each other by more than 2 kb

when positions were calculated with respect to mRNAs 5'ends, but less than 2 kb with respect to gene 5'ends or initiator ATGs in order to minimize the under-representation of conserved EREs due to differential annotation in the human and mouse genomes. For example, the EBAG9 gene element, which is perfectly conserved in sequence, would not be selected on the basis of the distances to the 5' end of mRNA because the mouse transcriptional start site was annotated at the ATG, but was included due to similar location with respect to the annotated 5'end of the genes. Note that the 5'end of the mRNA now coincides with the 5'end of the gene in the latest built of the mouse genome.

The probability of an ERE to be found by chance in both the human and mouse orthologs was calculated by multiplying the probability of finding an element within -10 to +5 kb of the start site of one gene (14,074 elements in that window of the human genome /34,699 total genes) by that of finding an ERE in a 4 kb window (+ or - 2kb) of corresponding sequence in the other species (12,828 elements x 4 kb /33,914 total mouse genes x 15 kb) and by the total number of orthologs (9,944 gene pairs) giving 407 expected elements.

Cell culture

Hela cells were maintained in DMEM (Wisent, St-Bruno, QC, Canada) supplemented with 5% fetal bovine serum (FBS, Sigma, Oakville, ONT, Canada). Cells were switched 3 days before experiments to medium without phenol red containing charcoal-stripped serum. For gel shift assays, Hela cells were electroporated (10^7 cells, 0.24 kV, 950 μ F in a Biorad Gene Pulser II apparatus) with 80 μ g expression vector (pSG5-ER α , pSG5-ER β or parental vector alone). Note that pSG5-ER β was generated by subcloning the open reading frame of hER β [72] from pCMVSPORT-ER β (a kind gift from Dr T. Willson, Glaxo, NC) into the BamH1 site of pSG5 [73] by PCR amplification. Cells were treated with 25 nM 17- β estradiol (Sigma) one to two hours before harvesting (48 h

post-transfection). Whole cells extracts were prepared by three freeze-thaw cycles in gel retardation buffer as previously described [74].

For ChIP experiments, MCF-7 and MDA-MB231 cells were grown in DMEM (Sigma) supplemented with 10% fetal calf serum (FCS, Sigma). The medium was changed to phenol-red free DMEM supplemented with 2.5% dextran-charcoal treated FCS 48 h before hormone addition, and replaced each 24 h. The MDA::hER α cell line stably expressing hER α was generated from hER α -negative MDA-MB231 cells by transfection of pCDNA3.1/Hygro-hER α [28].

Gel shift and ChIP assays

For gel shift assays, whole cells extracts expressing ER α or ER β or control extracts from cells transfected with the parental pSG5 vector were diluted to 120 mM KCl and assayed for binding to 32 P-labeled, double-stranded oligonucleotide probes (50,000 cpm/sample) as described before [74]. Radioactivity associated with bound or free probe was quantified using a Molecular Imager FX with the Quantity One software (BioRad, Mississauga, ONT, Canada).

For ChIP assays, chromatin was crosslinked using 1.5% formaldehyde for 5 min at 37°C and fragmented by sonication as previously reported [28, 36], yielding fragments of average size \sim 350 bp. Antibodies against a C-terminal epitope of hER α (HC20) and against TBP were purchased from Santa Cruz (Heidelberg, Germany), and the antibody against phosphorylated polymerase II was purchased from Upstate Biotechnology (Buckingham, England). The sequences of the primers used in ChIP assays (synthesized by MWG GmbH, Ebersberg, Germany) are available upon request.

ACKNOWLEDGMENTS

We thank Thomas Bédouin for assistance with initial identification of EREs, Robert Sladek, Amy Hauth and Franz Lang for helpful discussions, and Marie Pageau (Biochemistry, U. de Montréal) for help with access to computers. This work was supported by Canadian Institutes of Health Research (CIHR) and Natural Sciences and Engineering Research Council (NSERC) grants to S.M., funding from Genome Canada and Genome Quebec for the Steroid Atlas project to S.M. and J.H.W, and by European Molecular Biology Organization (EMBO) and European Molecular Biology Laboratory (EMBL) grants to F.G. V.B. acknowledges support from a CIHR Strategic Program to the Montreal Center for Experimental Therapeutics in Cancer and from a CIHR post-doctoral fellowship. J.D. is recipient of a studentship from the Fonds de Recherche en Santé du Québec (FRSQ), and R.M. was supported by a long-term fellowship from EMBO. J.H.W. and S.M. are chercheurs-boursier of the FRSQ, and SM holds the CIBC Breast Cancer Research Chair at Université de Montréal.

NOTE ADDED IN PROOF:

A database of estrogen responsive genes was published while this article was in revision (ERGDB, [75]). This database, compiled from a literature review of experimentally identified estrogen target genes, includes putative ERE sequences found in some of these genes through screening of their promoter-proximal sequences with Dragon ERE Finder v.2.0. Our database includes all ERGDB EREs that correspond to sequences used in our screen (although positions may vary, as our study used a later release of the human and mouse genomes), but also ERE sequences found beyond the cut-off distance used for the ERGDB (-4500 to +500 bp from transcriptional start sites), or located in genes not currently known to respond to estrogen.

TABLE LEGENDS:

Table 1: Estrogen response elements in the human genome and corresponding elements in the mouse genome.

The sequence of previously characterized consensus EREs or elements with one or 2 mismatches in the human genome are shown with the corresponding references. Position of the EREs with respect to the 5' end of mRNAs is as indicated in genomic data files, and may differ slightly from the position indicated in the references (in brackets). All other EREs in this table (selected examples from Table 3) were identified through our genomic screens (except for a few elements in mouse orthologs, identified by asterisks, which due to a higher number of mismatches were identified manually). References provided for novel EREs describe estrogenic regulation of the corresponding gene. ^a EFP is now called ZNF147 in the human genome database. ^b Trim25, the EFP/ZNF147 mouse homolog, corresponds to LOC217069. ^c Greb1, also referenced as Greb1-pending, corresponds to 5730583K22Rik.

Table 2: EREs with identical sequences in human and mouse orthologs.

Human and mouse EREs with identical sequences were selected from the compilation of EREs located in known human (Hs) and mouse (Mm) orthologs with less than 2 kb difference in either their respective distances to the beginning of genes, mRNA or coding sequences (see supplementary material for a full list). Elements that are fully conserved in sequence in the two orthologs were further analyzed to determine whether they are part of an annotated coding sequence (CDS, bottom) or not (top). Four additional elements were overlapping a CDS in at least one species (middle). Note that some elements are listed twice, as they are found in the vicinity of two genes, and are shown in the orientation corresponding to each gene. Number pairs identify identical EREs. Elements in bold are known EREs. * mouse gene names are similar to the human names except for NDRG2 and ZNF142, whose counterparts are Ndr2 and Zfp142, respectively.

Table 3: EREs found in upregulated estrogen target genes.

The listed EREs were compiled by comparison with upregulated E2 target genes identified in gene expression studies as cited. The positions given are distances to the most upstream mRNA 5'-end for each gene (negative distances correspond to elements located upstream of the 5', and conversely). ERE sequences in bold correspond to consensus elements. Underlined genes contain EREs that are conserved in mouse (see supplementary data) according to criteria defined in Materials and Methods.

FIGURE LEGENDS:

Fig. 1. *In vitro* binding of ERs to consensus and near-consensus EREs.

A. Consensus ERE with base numbering used in this study. **B.** Binding of human ER α or ER β to response elements derived from the vitellogenin A2 (VitA2) ERE carrying a single replacement in one of the two arms of the palindrome. **C.** Binding of human ER α or ER β to response elements derived from the VitA2 ERE carrying two symmetrical replacements. The percentage of probe found in a specific complex with human ER α or ER β was quantified by phosphorimager and is shown with the corresponding standard error of the means (SEM).

Fig. 2. Distribution of EREs in the human and mouse genomes.

A. Distribution of consensus EREs on human chromosomes. The percentage of nucleotides in individual chromosomes is plotted, along with the percentage of total response elements found per chromosome. **B.** Distribution of consensus EREs on mouse chromosomes. **C.** Distribution of EREs found between -10 to +5 kb of mRNA 5' ends. The ratio between the number of EREs found and the expected number assuming a random distribution of these elements in the -10 to +5 interval was calculated for consecutive 1 kb segments of flanking sequences in the human (dark bars) and mouse (grey bars) genomes. For the elements conserved between the two genomes (white bars), the distribution was characterized only

between -8 to +3 kb in order not to introduce a bias in the 2 kb constraint in relative distances between the two species.

Fig. 3. *In vitro* binding of ER α to EREs identified in this study.

A. Relative binding of hER α to known and newly characterized EREs in estrogen target genes. **B.** Binding of hER α to the novel EREs characterized through genomic screening. Binding is expressed as the percentage radioactivity of specifically bound probe versus bound plus free probe.

Fig. 4: *In vivo* binding of human ER α to EREs in ER-positive breast cancer cells.

Interaction of hER α with selected EREs in the absence or presence of estrogen was investigated by chromatin-immunoprecipitation assay in ER α -negative MDA-MB231 (MDA) cells, in ER α -positive MCF7 cells or in MDA-MB231 cells stably transfected with an expression vector for hER α (MDA::hER α) and treated or not with estradiol for 2 h. Fragments containing putative EREs were immunoprecipitated from formaldehyde-fixed chromatin preparation with antibodies against hER α , TATA-binding protein (TBP) or phosphorylated polymerase II (P-PolII), or with preimmune rabbit IgG. A separate fragment spanning the transcription start site was used for PCR amplification when the start site was situated at more than 400 bp of the ERE. Whenever available, experimentally-mapped transcription start sites were included in the +1 fragments. Experiments were performed 2-3 times and a representative result is shown.

SUPPLEMENTARY DATA:

Supplementary online Table 1: EREs found in both human and mouse orthologs.

List of EREs found in both human and mouse orthologs with less than a 2 kb difference in their relative distance to the 5' end of the gene, the mRNA or the coding sequence.

REFERENCES

1. **Couse JF, Korach KS** 1999 Estrogen receptor null mice: what have we learned and where will they lead us? *Endocrine Rev* 20:358-417
2. **McEwen BS, Alves SE** 1999 Estrogen actions in the central nervous system. *Endocrine Rev* 20:279-307
3. **Jordan VC** 2001 Estrogen, selective estrogen receptor modulation, and coronary heart disease: something or nothing. *J Natl Canc Inst* 93:2-4
4. **Nilsson S, Makela S, Treuter E, et al.** 2001 Mechanisms of estrogen action. *Physiol Rev* 81:1535-1565
5. **Green S, Chambon P** 1988 Nuclear receptors enhance our understanding of transcription regulation. *Trends Genet* 4:309-314
6. **Beato M, Herrlich P, Schütz G** 1995 Steroid hormone receptors: many actors in search of a plot. *Cell* 83:851-857
7. **Mangelsdorf DJ, Thummel C, Beato M, et al.** 1995 The nuclear receptor family: the second decade. *Cell* 83:835-839
8. **Chawla A, Repa JJ, Evans RM, Mangelsdorf DJ** 2001 Nuclear receptors and lipid physiology: opening the X-files. *Science* 294:1866-1870
9. **Robinson-Rechavi M, Escriva Garcia H, Laudet V** 2003 The nuclear receptor superfamily. *J Cell Sci* 116:585-6.
10. **Krust A, Green S, Argos P, et al.** 1986 The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. *EMBO J* 5:891-897
11. **Mader S, Chambon P, White JH** 1993 Defining a minimal estrogen receptor DNA binding domain. *Nucleic Acid Res* 21:1125-1132
12. **Kumar V, Chambon P** 1988 The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell* 55:145-156

13. **Fawell SE, Lees JA, White R, Parker MG** 1990 Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor. *Cell* 60:953-962
14. **Cowley SM, Hoare S, Mosselman S, Parker MG** 1997 Estrogen receptors alpha and beta form heterodimers on DNA. *J Biol Chem* 272:19858-19862
15. **Kato S, Tora L, Yamauchi J, Masushige S, Bellard M, Chambon P** 1992 A far upstream estrogen response element of the ovalbumin gene contains several half-palindromic 5'-TGACC-3' motifs acting synergistically. *Cell* 68:731-742
16. **Vanacker J-M, Petterson K, Gustafsson J-A, Laudet V** 1999 Transcriptional targets shared by estrogen receptor-related (ERRs) and estrogen receptor (ER) α , but not ER β . *EMBO J* 18:4270-4279
17. **Klinge CM** 2001 Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Res* 29:2905-2919
18. **Klein-Hitpass L, Schorpp M, Wagner U, Ryffel GU** 1986 An estrogen-responsive element derived from the 5' flanking region of the *Xenopus* Vitellogenin A2 gene functions in transfected human cells. *Cell* 46:1053-1061
19. **Sanchez R, Nguyen D, Rocha W, White JH, Mader S** 2002 Diversity in the mechanisms of gene regulation by estrogen receptors. *Bioessays* 24:244-54.
20. **Glass CK** 1994 Differential recognition of target genes by nuclear receptor monomers, dimers and heterodimers. *Endocrine Rev* 15:391-407
21. **Khorasanizadeh S, Rastinejad F** 2001 Nuclear-receptor interactions on DNA-response elements. *Trends Biochem Sci* 26:384-390
22. **Chen H, Lin RJ, Xie W, Wilpitz D, Evans RM** 1999 Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase. *Cell* 98:675-686
23. **Burakov D, Crofts LA, Chang CP, Freedman LP** 2002 Reciprocal recruitment of DRIP/mediator and p160 coactivator complexes in vivo by estrogen receptor. *J Biol Chem* 277:14359-14362

24. **Shang Y, Hu X, DiRenzo J, Lazar MA, Brown M** 2000 Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* 103:843-52.
25. **Reid G, Hubner MR, Metivier R, et al.** 2003 Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling. *Mol Cell* 11:695-707.
26. **Leo C, Chen JD** 2000 The SRC family of nuclear receptor coactivators. *Gene* 245:1-11
27. **Robyr D, Wolffe A, Wahli W** 2000 Nuclear hormone receptor coregulators in action: diversity for shared tasks. *Mol Endocrinol* 14:329-346.
28. **Rosenfeld MG, Glass CK** 2001 Coregulator codes of transcriptional regulation by nuclear receptors. *J. Biol. Chem.* 276:36865-68
29. **Rachez C, Freedman LP** 2001 Mediator complexes and transcription. *Curr Opin Cell Biol* 13:274-80.
30. **Dilworth FJ, Chambon P** 2001 Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene* 20:3047-3054
31. **McDonnell DP, Norris JD** 2002 Connections and regulation of the human estrogen receptor. *Science* 296:1642-1644
32. **Belandia B, Parker MG** 2003 Nuclear receptors: a rendezvous for chromatin remodeling factors. *Cell* 114:277-280
33. **Métivier R, Penot G, Hubner MR, et al.** 2003 Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 115:751-63
34. **Kushner PJ, Agard DA, Greene GL, et al.** 2000 Estrogen receptor pathways and AP-1. *J Steroid Biochem Mol Biol* 74:311-317
35. **Safe S** 2001 Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions. *Vitam. Horm.* 62:231-52
36. **Kelly MJ, Levin ER** 2001 Rapid actions of plasma membrane estrogen receptors. *Trends Endocrinol Metab* 12:152-156

37. **McEwen B, Akama K, Alves S, et al.** 2001 Tracking the estrogen receptor in neurons: implications for estrogen-induced synapse formation. *Proc Natl Acad Sci U S A* 98:7093-7100
38. **Moggs JG, Orphanides G** 2001 Estrogen receptors: orchestrators of pleiotropic cellular responses. *EMBO Rep* 2:775-781
39. **Inoue S, Orimo A, Hosoi T, et al.** 1993 Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *Proc Natl Acad Sci U S A* 90:11117-11121
40. **Watanabe T, Inoue S, Hiroi H, Orimo A, Kawashima H, Muramatsu M** 1998 Isolation of estrogen-responsive genes with a CpG island library. *Mol Cell Biol* 18:442-9.
41. **Ikeda K, Sato M, Tsutsumi O, et al.** 2000 Promoter analysis and chromosomal mapping of human EBAG9 gene. *Biochem Biophys Res Commun* 273:654-660
42. **Lander ES, Linton LM, Birren B, et al.** 2001 Initial sequencing and analysis of the human genome. *Nature* 409:860-921
43. **Tsai S-J, Wu M-H, Chen H-M, Chuang P-C, Wing L-YC** 2002 Fibroblast Growth Factor-9 Is an Endometrial Stromal Growth Factor. *Endocrinology* 143:2715-2721
44. **Cohen RI, Chandross KJ** 2000 Fibroblast growth factor-9 modulates the expression of myelin related proteins and multiple fibroblast growth factor receptors in developing oligodendrocytes. *J Neurosci Res* 61:273-287
45. **Colvin JS, Green RP, Schmahl J, Capel B, Ornitz DM** 2001 Male-to-female sex reversal in mice lacking fibroblast growth factor 9. *Cell* 104:875-889
46. **Toshima J, Ohashi K, Okano I, et al.** 1995 Identification and characterization of a novel protein kinase, TESK1, specifically expressed in testicular germ cells. *J Biol Chem* 270:31331-7.
47. **Toshima J, Koji T, Mizuno K** 1998 Stage-specific expression of testis-specific protein kinase 1 (TESK1) in rat spermatogenic cells. *Biochem Biophys Res Commun* 249:107-12.

48. **Venables JP, Vernet C, Chew SL, et al.** 1999 T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis. *Hum Mol Genet* 8:959-69.
49. **Valve E, Penttila TL, Paranko J, Harkonen P** 1997 FGF-8 is expressed during specific phases of rodent oocyte and spermatogonium development. *Biochem Biophys Res Commun* 232:173-7.
50. **Kang HY, Yeh S, Fujimoto N, Chang C** 1999 Cloning and characterization of human prostate coactivator ARA54, a novel protein that associates with the androgen receptor. *J Biol Chem* 274:8570-8576
51. **Couse JE, Mahato D, Eddy EM, Korach KS** 2001 Molecular mechanism of estrogen action in the male: insights from the estrogen receptor null mice. *Reprod Fertil Dev* 13:211-219
52. **Murata Y, Robertson KM, Jones ME, Simpson ER** 2002 Effect of estrogen deficiency in the male: the ArKO mouse model. *Mol Cell Endocrinol* 193:7-12
53. **Carreau S** 2003 Estrogens--male hormones? *Folia Histochem Cytobiol* 41:107-111
54. **Soulez M, Parker MG** 2001 Identification of novel oestrogen receptor target genes in human ZR75-1 breast cancer cells by expression profiling. *J Mol Endocrinol* 27:259-74.
55. **Mally MI, Cirulli V, Otonkoski T, Soto G, Hayek A** 1996 Ontogeny and tissue distribution of human GAD expression. *Diabetes* 45:496-501
56. **McCarthy MM, Kaufman LC, Brooks PJ, Pfaff DW, Schwartz-Giblin S** 1995 Estrogen modulation of mRNA levels for the two forms of glutamic acid decarboxylase (GAD) in female rat brain. *J Comp Neurol* 360:685-697
57. **Bosma PT, Blazquez M, Fraser EJ, Schulz RW, Docherty K, Trudeau VL** 2001 Sex steroid regulation of glutamate decarboxylase mRNA expression in goldfish brain is sexually dimorphic. *J Neurochem* 76:945-956
58. **Wondisford FE, Radovick S, Moates JM, Usala SJ, Weintraub BD** 1988 Isolation and characterization of the human thyrotropin beta-subunit gene.

- Differences in gene structure and promoter function from murine species. *J Biol Chem* 263:12538-42.
59. **Gurr JA, Vrontakis ME, Athanasian EA, Wagner CR, Kourides IA** 1986 Hormonal regulation of thyrotropin alpha and beta subunit mRNAs. *Horm Metab Res* 18:382-385
 60. **Podvinec M, Kaufmann MR, Handschin C, Meyer UA** 2002 NUBIScan, an in silico approach for prediction of nuclear receptor response elements. *Mol Endocrinol* 16:1269-79
 61. **Bajic VB, Tan SL, Chong A, et al.** 2003 Dragon ERE Finder version 2: a tool for accurate detection and analysis of estrogen response elements in vertebrate genomes. *Nucleic Acids Res* 31:3605-3607
 62. **Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M, Brown PO** 2001 Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409:533-538
 63. **Lieb JD, Liu X, Botstein D, Brown PO** 2001 Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet* 28:327-334
 64. **Topalidou I, Thireos G** 2003 Gcn4 occupancy of open reading frame regions results in the recruitment of chromatin-modifying complexes but not the mediator complex. *EMBO Rep* 4:872-6.
 65. **Martone R, Euskirchen G, Bertone P, et al.** 2003 Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci U S A* 100:12247-12252
 66. **Toft D, Gorski J** 1966 A receptor molecule for estrogens: isolation from the rat uterus and preliminary characterization. *Proc Natl Acad Sci U S A* 55:1574-81.
 67. **Clark JH, Gorski J** 1970 Ontogeny of the estrogen receptor during early uterine development. *Science* 169:76-78
 68. **Notides AC** 1970 The binding affinity and specificity of the estrogen receptor of the rat uterus and anterior pituitary. *Endocrinology* 87:987-92.

69. **Stumpf WE, Sar M, Zuber TJ, Soini E, Tuohimaa P** 1981 Quantitative assessment of steroid hormone binding sites by thaw-mount autoradiography. *J Histochem Cytochem* 29:201-6.
70. **Fernandez PC, Frank SR, Wang L, et al.** 2003 Genomic targets of the human c-Myc protein. *Genes Dev* 17:1115-1129
71. **Bhat RA, Harnish DC, Stevis PE, Lyttle CR, Komm BS** 1998 A novel human estrogen receptor beta: identification and functional analysis of additional N-terminal amino acids. *J Steroid Biochem Mol Biol* 67:233-240
72. **Green S, Issemann I, Sheer E** 1988 A versatile in vivo and in vitro eukaryotic expression vector for protein engineering. *Nucleic Acids Res* 16:369
73. **Nguyen D, Steinberg SV, Rouault E, et al.** 2001 A G577R mutation in the human AR P box results in selective decreases in DNA binding and in partial androgen insensitivity syndrome. *Mol Endocrinol* 15:1790-1802
74. **Agarwal A, Yeung WS, Lee KF** 2002 Cloning and characterization of the human oviduct-specific glycoprotein (HuOGP) gene promoter. *Mol Hum Reprod* 8:167-175
75. **Takahashi K, Sendai Y, Matsuda Y, Hoshi H, Hiroi M, Araki Y** 2000 Mouse Oviduct-Specific Glycoprotein Gene: Genomic Organization and Structure of the 5'-Flanking Regulatory Region. *Biol Reprod* 62:217-226
76. **Zhao YY, Zhou J, Narayanan CS, Cui Y, Kumar A** 1999 Role of C/A polymorphism at -20 on the expression of human angiotensinogen gene. *Hypertension* 33:108-15.
77. **Congiu M, Clouston WM, Fernley RT, Richards RI** 1992 The function of conserved elements in the promoter of the mouse angiotensinogen gene. *J Mol Endocrinol* 9:19-29
78. **Berry M, Nunez A-M, Chambon P** 1989 Estrogen-responsive element of the human pS2 gene is an imperfectly palindromic sequence. *Proc Natl Acad Sci USA* 86:1218-1222

79. **Richard S, Zingg HH** 1990 The human oxytocin promoter is regulated by estrogens. *J Biol Chem* 265:6098-6103
80. **Perillo B, Sasso A, Abbondanza C, Palumbo G** 2000 17beta -Estradiol Inhibits Apoptosis in MCF-7 Cells, Inducing bcl-2 Expression via Two Estrogen-Responsive Elements Present in the Coding Sequence. *Mol. Cell. Biol.* 20:2890-2901
81. **Citarella F, Misiti S, Felici A, Aiuti A, La Porta C, Fantoni A** 1993 The 5' sequence of human Factor XII gene contains transcription regulatory elements typical of liver specific, estrogen-modulated genes. *Biochim Biophys Acta* 1172:197-199
82. **Kyo S, Takakura M, Kanaya T, et al.** 1999 Estrogen Activates Telomerase. *Cancer Res* 59:5917-5921
83. **Fan JD, Wagner BL, McDonnell DP** 1996 Identification of the sequences within the human complement 3 promoter required for estrogen responsiveness provides insight into the mechanism of tamoxifen mixed agonist activity. *Mol Endocrinol* 10:1605-1616
84. **Augereau P, Miralles F, Cavailès V, Gaudet C, Parker M, Rochefort H** 1994 Characterization of the proximal estrogen-responsive element of human cathepsin D gene. *Mol Endocrinol* 8:693-703
85. **Teng CT, Liu Y, Yang N, Walmer D, Panella T** 1992 Differential molecular mechanism of the estrogen action that regulates lactoferrin gene in human and mouse. *Mol Endocrinol* 6:1969-1981
86. **Lee M, Liu Y, Zhang X** 1995 A retinoic acid response element that overlaps an estrogen response element mediates multihormonal sensitivity in transcriptional activation of the lactoferrin gene. *Mol Cell Biol* 15:4194-4207
87. **Lobenhofer EK, Bennett L, Cable PL, Li L, Bushel PR, Afshari CA** 2002 Regulation of DNA replication fork genes by 17beta-estradiol. *Mol Endocrinol* 16:1215-1229

88. **Revillion F, Pawlowski V, Hornez L, Peyrat JP** 2000 Glyceraldehyde-3-phosphate dehydrogenase gene expression in human breast cancer. *Eur J Cancer* 36:1038-42.
89. **Ghosh MG, Thompson DA, Weigel RJ** 2000 PDZK1 and GREB1 are estrogen-regulated genes expressed in hormone-responsive breast cancer. *Cancer Res* 60:6367-6375
90. **Hashimoto S, Miwa M, Akasofu K, Nishida E** 1991 Changes in 40 serum proteins of post-menopausal women. *Maturitas* 13:23-33
91. **Inoue A, Yoshida N, Omoto Y, et al.** 2002 Development of cDNA microarray for expression profiling of estrogen-responsive genes. *J Mol Endocrinol* 29:175-192
92. **Bhat KP, Pezzuto JM** 2001 Resveratrol exhibits cytostatic and antiestrogenic properties with human endometrial adenocarcinoma (Ishikawa) cells. *Cancer Res* 61:6137-6144
93. **Apparao KB, Lovely LP, Gui Y, Lininger RA, Lessey BA** 2002 Elevated endometrial androgen receptor expression in women with polycystic ovarian syndrome. *Biol Reprod* 66:297-304
94. **Pedram A, Razandi M, Aitkenhead M, Hughes CC, Levin ER** 2002 Integration of the non-genomic and genomic actions of estrogen. Membrane-initiated signaling by steroid to transcription and cell biology. *J Biol Chem* 277:50768-75
95. **Charpentier AH, Bednarek AK, Daniel RL, et al.** 2000 Effects of estrogen on global gene expression: identification of novel targets of estrogen action. *Cancer Res* 60:5977-83
96. **Xie T, Ho SL, Ramsden D** 1999 Characterization and implications of estrogenic down-regulation of human catechol-O-methyltransferase gene transcription. *Mol Pharmacol* 56:31-8
97. **Seth P, Krop I, Porter D, Polyak K** 2002 Novel estrogen and tamoxifen induced genes identified by SAGE (Serial Analysis of Gene Expression). *Oncogene* 21:836-43

98. **Zhang SY, Liu SC, Al-Saleem LF, et al.** 2000 E2F-1: a proliferative marker of breast neoplasia. *Cancer Epidemiol Biomarkers Prev* 9:395-401
99. **Seo HS, Larsimont D, Querton G, et al.** 1998 Estrogenic and anti-estrogenic regulation of estrogen receptor in MCF-7 breast-cancer cells: comparison of immunocytochemical data with biochemical measurements. *Int J Cancer* 78:760-5
100. **Shimomura Y, Matsuo H, Samoto T, Maruo T** 1998 Up-regulation by progesterone of proliferating cell nuclear antigen and epidermal growth factor expression in human uterine leiomyoma. *J Clin Endocrinol Metab* 83:2192-8
101. **Kuang WW, Thompson DA, Hoch RV, Weigel RJ** 1998 Differential screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor-positive breast carcinoma cell line. *Nucleic Acids Res* 26:1116-1123
102. **Piotrowicz RS, Ding L, Maher P, Levin EG** 2001 Inhibition of cell migration by 24-kDa fibroblast growth factor-2 is dependent upon the estrogen receptor. *J Biol Chem* 276:3963-70
103. **Kumar P, Mark PJ, Ward BK, Minchin RF, Ratajczak T** 2001 Estradiol-regulated expression of the immunophilins cyclophilin 40 and FKBP52 in MCF-7 breast cancer cells. *Biochem Biophys Res Commun* 284:219-225
104. **Duan R, Porter W, Safe S, Weisz A, Rosales R** 1998 Estrogen-induced c-fos protooncogene expression in MCF-7 human breast cancer cells: role of estrogen receptor Sp1 complex formation. *Endocrinology* 139:1981-1990
105. **Daly RJ, Gu H, Parmar J, et al.** 2002 The docking protein Gab2 is overexpressed and estrogen regulated in human breast cancer. *Oncogene* 21:5175-5181
106. **Ormandy CJ, Lee CS, Ormandy HF, et al.** 1998 Amplification, expression, and steroid regulation of the preprogalanin gene in human breast cancer. *Cancer Res* 58:1353-7.
107. **Kang SK, Choi KC, Tai CJ, Auersperg N, Leung PC** 2001 Estradiol regulates gonadotropin-releasing hormone (GnRH) and its receptor gene expression and

antagonizes the growth inhibitory effects of GnRH in human ovarian surface epithelial and ovarian cancer cells. *Endocrinology* 142:580-588

108. **Inadera H, Hashimoto S, Dong HY, et al.** 2000 WISP-2 as a novel estrogen-responsive gene in human breast cancer cells. *Biochem Biophys Res Commun* 275:108-114
109. **Tschugguel W, Stonek F, Zhegu Z, et al.** 2001 Estrogen increases endothelial carbon monoxide, heme oxygenase 2, and carbon monoxide-derived cGMP by a receptor-mediated system. *J Clin Endocrinol Metab* 86:3833-9.
110. **Lee AV, Jackson JG, Gooch JL, et al.** 1999 Enhancement of insulin-like growth factor signaling in human breast cancer: estrogen regulation of insulin receptor substrate-1 expression in vitro and in vivo. *Mol Endocrinol* 13:787-796
111. **Molloy CA, May FE, Westley BR** 2000 Insulin receptor substrate-1 expression is regulated by estrogen in the MCF-7 human breast cancer cell line. *J Biol Chem* 275:12565-12571
112. **Mauro L, Salerno M, Panno ML, et al.** 2001 Estradiol increases IRS-1 gene expression and insulin signaling in breast cancer cells. *Biochem Biophys Res Commun* 288:685-689
113. **Oesterreich S, Zhang P, Guler RL, et al.** 2001 Re-expression of estrogen receptor alpha in estrogen receptor alpha-negative MCF-7 cells restores both estrogen and insulin-like growth factor-mediated signaling and growth. *Cancer Res* 61:5771-7.
114. **Pentecost BT** 1998 Expression and estrogen regulation of the HEM45 mRNA in human tumor lines and in the rat uterus. *J Steroid Biochem Mol Biol* 64:25-33
115. **Choi I, Gudas LJ, Katzenellenbogen BS** 2000 Regulation of keratin 19 gene expression by estrogen in human breast cancer cells and identification of the estrogen responsive gene region. *Mol Cell Endocrinol* 164:225-37
116. **Seth P, Porter D, Lahti-Domenici J, Geng Y, Richardson A, Polyak K** 2002 Cellular and molecular targets of estrogen in normal human breast tissue. *Cancer Res* 62:4540-4.

117. **Moalli PA, Klingensmith WL, Meyn LA, Zyczynski HM** 2002 Regulation of matrix metalloproteinase expression by estrogen in fibroblasts that are derived from the pelvic floor. *Am J Obstet Gynecol* 187:72-79
118. **Harendza S, Lovett DH, Panzer U, Lukacs Z, Kuhn P, Stahl RA** 2003 Linked common polymorphisms in the gelatinase a promoter are associated with diminished transcriptional response to estrogen and genetic fitness. *J Biol Chem* 278:20490-20499
119. **Marin-Castano ME, Elliot SJ, Potier M, et al.** 2003 Regulation of estrogen receptors and MMP-2 expression by estrogens in human retinal pigment epithelium. *Invest Ophthalmol Vis Sci* 44:50-59
120. **Kleinert H, Wallerath T, Euchenhofer C, Ihrig-Biedert I, Li H, Forstermann U** 1998 Estrogens increase transcription of the human endothelial NO synthase gene: analysis of the transcription factors involved. *Hypertension* 31:582-588
121. **Skaar TC, Prasad SC, Sharareh S, Lippman ME, Brunner N, Clarke R** 1998 Two-dimensional gel electrophoresis analyses identify nucleophosmin as an estrogen regulated protein associated with acquired estrogen-independence in human breast cancer cells. *J Steroid Biochem Mol Biol* 67:391-402.
122. **Prange-Kiel J, Rune GM, Zwirner M, Wallwiener D, Kiesel L** 2001 Regulation of estrogen receptor alpha and progesterone receptor (isoform A and B) expression in cultured human endometrial cells. *Exp Clin Endocrinol Diabetes* 109:231-7
123. **Martini PG, Katzenellenbogen BS** 2001 Regulation of prothymosin alpha gene expression by estrogen in estrogen receptor-containing breast cancer cells via upstream half-palindromic estrogen response element motifs. *Endocrinology* 142:3493-5501
124. **Wang Z, Kyo S, Maida Y, et al.** 2002 Tamoxifen regulates human telomerase reverse transcriptase (hTERT) gene expression differently in breast and endometrial cancer cells. *Oncogene* 21:3517-24.

125. **Gruvberger S, Ringner M, Chen Y, et al.** 2001 Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* 61:5979-5984
126. **Qin C, Nguyen T, Stewart J, Samudio I, Burghardt R, Safe S** 2002 Estrogen up-regulation of p53 gene expression in MCF-7 breast cancer cells is mediated by calmodulin kinase IV-dependent activation of a nuclear factor kappaB/CCAAT-binding transcription factor-1 complex. *Mol Endocrinol* 16:1793-809
127. **Duque G, Abdaimi KE, Macoritto M, Miller MM, Kremer R** 2002 Estrogens (E2) regulate expression and response of 1,25-dihydroxyvitamin D3 receptors in bone cells: changes with aging and hormone deprivation. *Biochem Biophys Res Commun* 299:446-454
128. **Katoh M** 2003 Expression and regulation of WNT1 in human cancer: up-regulation of WNT1 by beta-estradiol in MCF-7 cells. *Int J Oncol* 22:209-212
129. **Kirikoshi H, Katoh M** 2002 Expression and regulation of WNT10B in human cancer: up-regulation of WNT10B in MCF-7 cells by beta-estradiol and down-regulation of WNT10B in NT2 cells by retinoic acid. *Int J Mol Med* 10:507-511
130. **Kirikoshi H, Sekihara H, Katoh M** 2001 Expression of WNT14 and WNT14B mRNAs in human cancer, up-regulation of WNT14 by IFNgamma and up-regulation of WNT14B by beta-estradiol. *Int J Oncol* 19:1221-1225
131. **Levenson AS, Svoboda KM, Pease KM, et al.** 2002 Gene expression profiles with activation of the estrogen receptor alpha-selective estrogen receptor modulator complex in breast cancer cells expressing wild-type estrogen receptor. *Cancer Res* 62:4419-4426
132. **Frasor J, Danes JM, Komm B, Chang KC, Lyttle CR, Katzenellenbogen BS** 2003 Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology* 144:4562-4574
133. **Monroe DG, Getz BJ, Johnsen SA, Riggs BL, Khosla S, Spelsberg TC** 2003 Estrogen receptor isoform-specific regulation of endogenous gene expression in

human osteoblastic cell lines expressing either ERalpha or ERbeta. J Cell Biochem 90:315-326

134. **Tang S, Han H, Bajic VB** 2004 ERGDB: Estrogen Responsive Genes Database. Nucleic Acids Res 32 Database issue:D533-6

Table 1

HUMAN GENOME					MOUSE GENOME				
Symbol	UniGene#	ERE sequence	Pos.	Ref	Symbol	UniGene#	ERE sequence	Pos.	Ref
PREVIOUSLY CHARACTERIZED EREs									
Perfect elements									
EBF ^a	Hs 1579	GGGTCATGGTGACCC	+23,088	39	Trm25 ^b	Mm.248445	AGGqCAGGGTGACCT	+17,499	
EBAG9	Hs 9222	GGGTCAGGGTGACCT	+6 (-64)	40	Ebag9	Mm.287896	GGGTCAGGGTGACCT	-4886	
COX7A2L	Hs 423404	GGGTCAAGGTGACCC	+383 (+443)	40	Cox7A2l	Mm.30072	GGGTCACTTGACCC	+336	
1 mismatch									
OVGP1	Hs 1154	GGGTCACTGTGACtC	-156 (-170)	74	Ovgp1	Mm.5110	cGGTCATTGTGACtC	-110 (-110)	75
AGT	Hs 19383	AGGqCATCGTGACCC	-10 (-25)	76	Agt	Mm.285467	AGGTCACTTGACCC	-2379 (-2742)	77
2 mismatches									
pS2 / TFF1	Hs 350470	AGGTCAAGGTGgCCa	-392 (-406)	78	Tff1	Mm.2854	ctGTCACTCTTgCCa	-462*	
OXT	Hs 113216	cGGTgACCTTGACCC	-166 (-168)	79	Oxt	Mm.16745	cGaTgACCTTGACCC	-210*	
BCL2	Hs 79241	tGGTCcACCTGACCC	+319 (+276)	80	Bcl2	Mm.233518	tGGTCcATCTGACCC	+279	
FI2	Hs 1321	AGGqCAGCTTGACCa	-30 (-47)	81	FI2	Mm.42224	GaGcAAGCTTGACCa	-75*	
TERT	Hs 439911	tGGTCAGGCTGAtCT	-2687 (-2677)	82	Tert	Mm.10109			
C3	Hs 284394	AGGTgGCCCTGACCC	-226 (-237)	83	C3	Mm.19131	AtcTggCCTTGACCC	-208*	
CTSD	Hs 343475	GGGcCqGGCTGACCC	-125 (-270)	84	Ctd	Mm.231395			
LTF	Hs 437457	AGGTCAAGGcGAtCT	-81 (-358)	85	Ltf	Mm.282359	AGGTCAAGGTaACCC	-340 (-342)	86
OTHER ELEMENTS FOUND IN ESTROGEN TARGET GENES WITH KNOWN EREs									
C3	Hs 284394	GGGTcTtTGTGACCG	-9210						
CTSD	Hs 343475	GGGcCACCATGACCC	-8763		Ctd	Mm.231395	AGGcCAATCTGACCT	-8157	
TERT	Hs 439911	GGGcCAGAGTGACCC	-5663						
OXT	Hs 113216	AGGTCAAGCTGACCG	-2465		Oxt	Mm.16745	tGGTCACCGTGAtCC	-5424	
LTF	Hs 437457	AGGTcAGTCTGACCa	-1166						
OVGP1	Hs 1154	GGGTCCCTCTGACCT	+200		Ovgp1	Mm.5110	AGGTCAAGATGtCtg	+866*	
COX7A2L	Hs 423404	AGGTCAAGATGtCCa	-5246						
EBF ^a	Hs 1579	AGGqCAGGGTGACCT	+22,583		Trm25 ^b	Mm.248445	AGGqCAGGGTGACCT	+17,499	
SELECTED NEW ELEMENTS (used for gel shift and/or ChIP experiments)									
ADORA1	Hs 77867	AGGTtAGGGTGACCT	-1171	54					
		GGGTCAAGGTGAACT	-486						
		GGGTcGCGGTGAACT	-392						
		GGGTCAAGGTGAACT	-312						
CASP7	Hs 9216	GGGTCAAGGTGAACT	-272	87	Casp7	Mm.298737	GtGTCAACTTGACCa	+1458	
		GGGTCAAGGTGAAcG	-232						
		tGGTCAGGGTGAACT	-195						
		GGGTCAAGGTGAACT	-141						
GAPD	Hs 169476	AGGcCATCGTGACCT	-738	88	Gapd	Mm.288146	AGGTCAAGATGcCCT	-745	
		AGGTCAAAATGACCT	-21,207						
GREB1	Hs 438037	AGGTCAATCATGACCT	-9517	89	Grebl ^c	Mm.218957	AGGTcGCGTGACCC	-7666	
		GGGTCAATCTTGACCT	-1583				AGGTCAAGATGACCC	-3289	
IGFBP4	Hs 1516	AGGTCAATGTGACaC	-4124	54	Igfbp4	Mm.233799	AGaTCACCGTGACCT	-6697	
LY6E	Hs 77667	GGGcCAAGATGACCT	-499	54					
NR1P1	Hs 155017	GGGTCACTTTGACCC	-706	54	Nrp1	Mm.20895	AGGTCAATTTTGACCC	-636	
SCNN1A	Hs 446415	AGGTCAAGCTcACCC	-559	54					
GAD2	Hs 231829	AGGTcGCACTGACCT	-564		Gad2	Mm.4784	AGGTCAAGcGACCT	-349	
TSHB	Hs 406687	AGGTCAAGCTTGACaT	-4099	59	Tshb	Mm.110730	AtGTCAAACTGACCT	-535	

Table 2

Human Gene Name ^a	UniGene #	ERE Sequence	Hs Dist. ERE-mRNA	Mm Dist. ERE-mRNA	Dist. Dist. Mm-Hs ERE-Genes	Dist. Dist. Mm-Hs ERE-mRNA	Dist. Dist. Mm-Hs ERE-CDS
EREs located in a non-coding sequence (CDS)							
HOXD13	Hs.152414	tGGTCAAAATGACCC	-8865	-8719	146	146	257
KHDRBS3	Hs.13565	AGGTCATGATGcCCA	-7491	-6925	552	566	724
LDB1	Hs.26002	GGGTCACTTTGACaC	-5679	-5682	5869	3	520
RNF14 ¹	Hs.170926	AGcTCAAGTTGACCC	-5410	-6671	3522	1261	532
HOXC12	Hs.381267	AGGTCTTTCTGACCT	-5201	-5113	88	88	88
PCDH12 ¹	Hs.115897	GGGTCAACTTGAgCT	-4640	-5554	872	914	10
KIF24 ¹	Hs.201179	tGGTCAGACTGACaT	-2923	-3246	323	323	7425
PDE8A	Hs.78746	cTGTCAGTGTGACCC	-2583	-1914	799	669	674
NDRG2	Hs.243960	GaGTCACTCTGACCC	-1347	-1900	178	553	21
TESK1	Hs.79358	GGGTCAACAATGACTa	-991	-981	1457	10	260
NKX6-1	Hs.347185	GGGTCAAGTTTGCCa	-963	-410	353	553	82
PHF5A ¹	Hs.273234	GGGaCgCCTTGACCT	-659	-807	148	148	153
SLC10A2	Hs.194783	cTGTCACCTTGACCT	-382	-753	542	371	177
FGF9	Hs.111	tGGTCATTCTGATcT	-204	-237	2338	33	6
EBA9	Hs.9222	GGGTCAAGGTGACCT	6	-4886	226	4892	5816
FGF8	Hs.57710	GGGaCgGCGTGACCC	67	153	128	86	67
CNNM4	Hs.175043	tGGTCAACGTGAgCC	215	57	158	158	42
ACO2 ³	Hs.300463	AGGTCAAGGcGtCCC	219	203	122	16	2
COL4A2	Hs.407912	AGGTgAGAGcGACCC	243	284	41	41	26
EDG2	Hs.75794	AGGTCAAGTgGcCCC	257	432	518	175	16676
COX7A2L	Hs.423404	GGGTCAAGGTGACCC	381	336	7812	45	16
UBE2R2	Hs.11184	AGGgCAGGGTGACCT	755	715	279	40	20
NFKBIA	Hs.81328	AGGTCTTTATGACCC	1233	1107	29	126	89
NR1D1	Hs.276916	GGtTCACCCCTGACC	2314	1673	88	641	89
NUDT2 ²	Hs.429142	AtGTCAAGCTGACCa	2617	2981	364	364	2470
HNRPDL	Hs.372673	GtGTCAACCTGACCa	4639	3226	518	1413	957
EREs located partially in a coding sequence (CDS)							
B4GALT7	Hs.54702	AGGTCACTCTGgCCC	-6523	-5409	1105	1114	1122
SIRT7 ⁴	Hs.184447	GGGTtATGAcGACCC	-5012	-4845	378	167	2786
SLC25A12	Hs.353282	cGGTCAGAGTGACTc	27	77	216	50	0
CD7	Hs.36972	AGGTCAAGTGTGAgCC	1313	1751	357	438	385
MAFG ⁴	Hs.252229	GGGTtATGAcGACCC	4516	9	969	4507	0
EREs located in a coding sequence							
DEF6	Hs.15476	AGGTCAAGCCTGtCCC	-4108	-4746	652	638	651
ARHGAP4 ⁵	Hs.3109	AGGTCAAGCGaGACC	-3795	-5798	1885	2003	1961
ZNF142 ⁶	Hs.80526	cGGTCAGCCaGACCC	-2345	-2383	5	38	728
GPR73L1	Hs.375029	tGGTgATTATGACCT	103	94	3818	9	9
ENDOG	Hs.420106	GGGcCgGCCTGACCC	170	64	473	34	140
EDN3	Hs.1408	GGtTCACAGTGACCT	220	376	333	106	0
SFRS2	Hs.73965	tGGaCAACCTGACCT	221	219	595	2	0
RARG	Hs.1497	GGGcCAGCCTGACCT	288	4880	232	4592	0
TPBG	Hs.82128	tGGTgAGCCTGACCT	820	2324	448	1504	18
FZD6	Hs.114218	GGGTCAATTATGACC	1355	270	176	1085	0
PAX9	Hs.132576	GGGTCAATTAcGACTc	1603	1309	267	294	94
AMBP	Hs.76177	AGaTCAGCATGACC	1692	1545	147	147	140
BCSL1 ⁶	Hs.150922	GGGTCTGGGTGACCg	1847	2020	448	173	65
PRSS8	Hs.75799	AGGTCAAGCATcACCT	2528	2209	268	319	199
APIG2	Hs.343244	tGGcCAACCTGACCC	4610	3756	474	854	98
ARD1 ⁵	Hs.433291	AGGTCAAGCGaGACC	4961	4916	122	45	100

Table 3

Gene	UniGene #	Pos.	ERE Sequence	REP	Gene	UniGene #	Pos.	ERE Sequence	REP	Gene	UniGene #	Pos.	ERE Sequence	REP
ADCC9	H:2086	-388	tGGTCACATgACCT	132	E2C5	H:3943	-8234	AGaTCAGCCTGACC	95	KRT7	H:2381	2161	GGGCAATGTGACCC	116
ADP	H:3416	147	AGGTCAAGTGTGACC	132	EBADP	H:5022	6	GGGTCAAGGTGACCT	40			2259	tGGTCATTTGTGAGCC	
ADORA1	H:7787	-3804	AGGTCAAGATGACCT	54	EDEM	H:15438	-7838	tGGTCAGTGTGACC	99	KRTHE3	H:18205	-8886	GGGTCAAGAGGACC	116
ADT	H:1938	-10	AGGTCAAGGTGACCT		EDCE	H:7594	-7418	ctGTCAACCTGACCC	133	LMBTL	H:30383	3533	GGGTCTGTGTGACC	94
AHD	H:20446	-3825	GGGTCAAGATGACCT	133			237	GGGgcGACGTGACCT		LCHA	H:2795	-1074	AGaTCAGCCTGACC	95
AM1	H:42380	369	GGGTCAAGCTGACCC	91	ECFR	H:77482	3465	GGaTCAGCTTGAACCT	100			325	AGGTCTGCTGACCT	
AKR1C4	H:17787	1657	GGGTCAATTTGACC	91	ECLE2	H:34677	-8459	GGaTCAGCTTGAACCT	97			381	AGGTggCAGTGACCC	
ALP1	H:5700	-5820	tGGTCACCTGgACCT	133	ECR3	H:7408	2355	AGGTCTGTGTGACCT	95	LOC9102	H:43843	-5139	tGGTCATATGACCT	54
		-419	AGGTCTGGCTGACCC		EFSA	H:33021	1913	AGGTCTGTGTGACCT				882	AGGTggCAGTGACCC	
		2170	AGGacCAAGTGAACCC				3265	AGGTCAATTTGAGCC		LTF	H:43845	-1165	AGGTgAGTGTGACC	85
ALP	H:23425	-4383	GGaTCACAGGACCC	92	ELAZA	H:21	-7130	AGGTCAAGTGAACCC	133			-81	AGATCTGCTGACCT	
AR	H:59915	-559	GGGTgAAATGACCT	99	EPO	H:2333	-1654	AGGTCAATTTGACCT	94	LYE	H:7767	-9749	AGGTCAAGCTGgGCC	54
ARHC	H:19795	-5916	AGGTCAAGGACCT	91	F2	H:131	-30	GGGacCAAGTGAACCT	21			-469	GGGacAAGTGAACCT	
		-5815	AGGacCAAGTGAACCC		FBP1	H:30359	4539	AGGTCAATTTGACCT	91			358	cGGTCAACCCGACCC	
		-3815	GGaTCACAGTGAACCC		FBP2	H:30359	12	AGGTCAAGGTGAGC	91	MAPB	H:10342	-2800	AGGTgAAATGACC	91
ART3	H:24075	-5819	GGGTCAAGGACCT	94	FCR	H:23644	-3172	AGGTCAAGTGAACCC	102	MATN1	H:13365	-4045	GGGTCAAGGACCT	134
AS5	H:16785	365	GGGTCAAGTGTGACCT	94	FCR2	H:23644	-4294	AGGTggGCTGACCC	102	MCD307	H:16132	-731	AGGTCAAGGTGACCT	54
ATX32	H:8512	-3810	GGGTCAATTTGACCT	94	FCR3	H:111	204	AGGTCAATTTGACCT	94	MED	H:41975	515	AGGTCAAGGTGACCC	133
		388	GGGTCAAGGTGACCC		FR4	H:19475	85	AGGTCAATTTGACCT	94	MGP2	H:165	-553	AGGTCAAGGTGACCC	133
AVD1	H:52745	-4019	tGGTCAGAGGACCC	54	FRB4	H:848	105	AGGTCAATTTGACCT	95			-355	tGGTCATTTGACCT	
		-5940	AGGTCAAGTGTGACCTg		FCB	H:2547	-3411	AGGTCAAGCTGACCT	94	MM2E	H:36377	-2578	AGGTCTTTATGACCT	117
B4CAL1	H:36768	-3138	AGGTgAAATGACCC	54	FCB2	H:2547	-7848	tGGTCAGTGTGACCT	91			801	GGGTCTTTGTGACCT	118
		3042	AGGTCAAGCTGACCC		FCB3	H:17013	391	cGGTCAATTTGACCC	103			2255	GGGTCTTTGTGACCT	119
BDRB2	H:20382	-3398	tGGTCATGCTGACCT	133	CFB2	H:8305	-59	GGGTCAAGCTGACCC	54			3822	AGGTgATCTGACCT	
BCL2	H:7594	319	GGGTCAAGTGTGACCC	80			4555	AGGTCAATTTGACCT		MD2	H:8940	-6830	AGGTCAAGATGACCT	54
BCL2L1	H:8403	-4300	GGGTgATCTGACCT	94	QAB2	H:3087	-3889	AGGTCAATTTGACCT	105	MTFCE	H:15872	-4941	GGaTCACCTGACCT	91
BCL2L1	H:8403	529	AGGTCAAGCTGACCT	94	QAL	H:23859	-4992	tGGTCAGCTGACCC	106	MYB2	H:17718	3441	AGGTggCAAGTGAACCT	132
BIRC3	H:12789	-6855	AGGTgAAATGACCT	133			-4635	GGGTCAAGTGTGACCC		MYB3	H:17874	-2740	GGGTCAAGATGACCT	94
BIRC5	H:1573	-4882	tGGTCAGTGTGACCT	132	QATD	H:19465	-4881	AGGTCAATTTGACCT	94	MYB3	H:17874	-2740	tGGTCAGATGACCT	95
		4886	GGaTCACCTGACCT				-5970	GGGTCAATTTGACCT		NCB	H:44533	-1591	AGGTCAAGGCTGACCC	133
BTEB1	H:13757	-4883	AGGTCAATTTGACCT	132			-738	AGGTCAATTTGACCT				-4818	cGGTCAAGCTGACCC	
BUB1	H:28742	-4300	AGGTCAATTTGACCT	132	QAB3	H:49770	1780	GGGTCAAGCTGACCC	103	NM1	H:41038	-1576	AGGTCAAGCTGACCC	121
		-3880	AGGTCAAGCTGACCC		QCH1	H:8524	-3702	cGGTCAATTTGACCC	103	NR2	H:7838	-4535	AGGTCTGTGTGACCT	133
C11orf8	H:43010	-4309	AGGTgATTTGACCC	133			3748	AGGTCAAGATGACCC				-2519	AGGTCAATTTGACCT	
C3	H:28494	4210	GGGTCTTTGTGACCCg	133	CHRR	H:767	3749	tGGTCAGCTGACCT	103			705	GGGTCTGTGTGACCT	
CALCR	H:60	927	GGGTCAAGCTGACCC	132			4741	AGGTCAATTTGACCT		NR4A1	H:1119	3478	tGTCTACCTGACCT	94
CALM2	H:42818	-8818	cGGTCAATTTGACCT	95	CHRR	H:767	1619	tGGTCAGCTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CAS9	H:9216	157	AGGTCAAGCTGACCT	95	CHRR	H:767	3749	GGGTCAATTTGACCT	103			3743	GGGTCTGTGTGACCT	
		-382	AGGTCAAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
		312	GGGTCAAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
		272	GGGTCAAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
		232	GGGTCAAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
		195	tGGTCAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
		141	GGGTCAAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CAS8	H:20491	1374	AGGTCAATTTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CAN1	H:7034	-6950	tGGTCAGCTGACCT	95	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
		2616	GGGTCAAGCTGACCT		CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CBFA2T3	H:10039	-3378	GGGTCAAGCTGACCT	132	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCBR1	H:34265	1339	AGGTCTTTATGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB3	H:19781	82	GGGTCAAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB4	H:34690	-6306	GGGTCAAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB5	H:19401	-4894	GGGTCAAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB6	H:27954	-7718	GGGTCAAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB7	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB8	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB9	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB10	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB11	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB12	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB13	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB14	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB15	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB16	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB17	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB18	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB19	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB20	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB21	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB22	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB23	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB24	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB25	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB26	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB27	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB28	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB29	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB30	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB31	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB32	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB33	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB34	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB35	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB36	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB37	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB38	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT	103	NR4A3	H:27922	-1589	GGGTCAAGCTGACCC	94
CCB39	H:40681	-6134	tGGTCAGCTGACCT	133	CHRR	H:767	3749	GGGTCAATTTGACCT						

Table 3-suite

Gene	UniGene #	Pos	ERESequence	REF	Gene	UniGene #	Pos	ERESequence	REF	Gene	UniGene #	Pos	ERESequence	REF
SLC29A1	Hs25493	4242	cgggtcaccgttgacct	54	TERT	Hs439911	-5663	gggtcacccttgacc	124	VDR	Hs2062	-3231	gggtcaccctttacc	127
SLC7A5	Hs18401	4863	aaattcagaaatgacct	5491			-2687	agattcacccttgacc		VEGF	Hs73793	-9867	aggtcaccctttgacc	132
SLK	Hs105751	-3049	aggtcaccctgacc	132	TFPI	Hs330470	-392	tggcacccttgacct	91,133	VTG	Hs2257	-4039	aggtcaccatgacct	87
SMC2L1	Hs115023	-9528	aggtcctttttgacct	94	TFPI3	Hs82561	-8555	aggtcctttttgacct	54,124	VWF	Hs440848	-9547	aggtcaccagcggacc	133
		-8578	tgcctcacccttgacct				-4315	gaattcattcaccacct	125	WISP2	Hs194679	-2506	aaattcggatgacct	97
SNRK	Hs79025	-1342	aggtcaccagatgacc	94	TFPI2	Hs438231	-4126	tgggtcatttttacct	54			-402	gggtcaccaccacct	
SNRPA	Hs173255	2000	aggtcggcagatgacct	54	THBD	Hs2000	-2152	gggtgggacatgacc	94	WNT1	Hs248164	-9908	gggtcaccatgacct	128
SOS1	Hs336392	-3457	tggcctttttgacct	94	TNFRSF8	Hs1314	-6902	aggtcaccagatgacct	94	WNT10B	Hs91985	3215	gggtcaccatgacct	129
SP4	Hs2982	-1480	gggtcacccttcacct	94			-2559	gggtgaccttggacc		WNT9B	Hs326420	1839	gggtcaccctgacc	131
SPRY1	Hs436944	136	gggtcaccaccgacct	94	TDB1	Hs178137	-3159	aggtcacccttgacc	54			3736	gggtcggcctgacc	
SPRY2	Hs18576	1041	gggtcaccaccgacct	94	TOPBP1	Hs91417	-1262	tgaatcacccttgacct	133	XPNPEP1	Hs360629	-799	aggttggcattgacct	94
SSNA1	Hs18328	-8128	gggtcggatgacct	107	TP53	Hs408312	-1529	aggtcggatgacct	126	XROCA	Hs19090	-7225	aggttacccttgacc	133
		-1829	cgggtcaccctgacc		TPBG	Hs82128	-4705	tctgtcaccatgacc	54	ZDRHCA	Hs3298	-9951	gggtcaccctttacc	97
		3468	aggtcattctgacct				820	tgggtaccctgacct		ZNF17	Hs185796	3231	gggtcaccctgacct	94
STAG1	Hs138263	-4456	caggtcaccatgacct	133	TPO	Hs71304	-619	caggtcaccctgacct	94	ZNF230	Hs192983	-9478	gggtcaccaccgacct	94
STK6	Hs25822	3170	gggtcaccattccacct	132	TPX1	Hs2042	-7640	gggtcacccttgacc	133			82	gggtcggcaccgacct	
SUOX	Hs16340	-7469	aggtcaccatgacct	95			-3600	aggtcctttatgacct		ZNF75	Hs332015	-6059	aggtcaccatgacct	94
		305	aaattcaccctgacct		TRIM31	Hs49275	-288	aggtcaccatgacct	133			-62	gggtcaccattgacct	
		1718	gggtcaccggatgacct		TSHB	Hs406287	-4099	aggtcacccttgacct	99	ZNF9	Hs2110	2272	aggtcggcattgacct	95
TAC1	Hs2563	167	gggtcaccaccgacct	94	USP2	Hs454534	305	gggtcggcctgacct	97	ZNF91	Hs8297	-5023	aaattcaccatgacct	131
TCF2	Hs408093	996	aggtcaccattgacct	94			2061	gggtcaccctgacct						

Figure 1

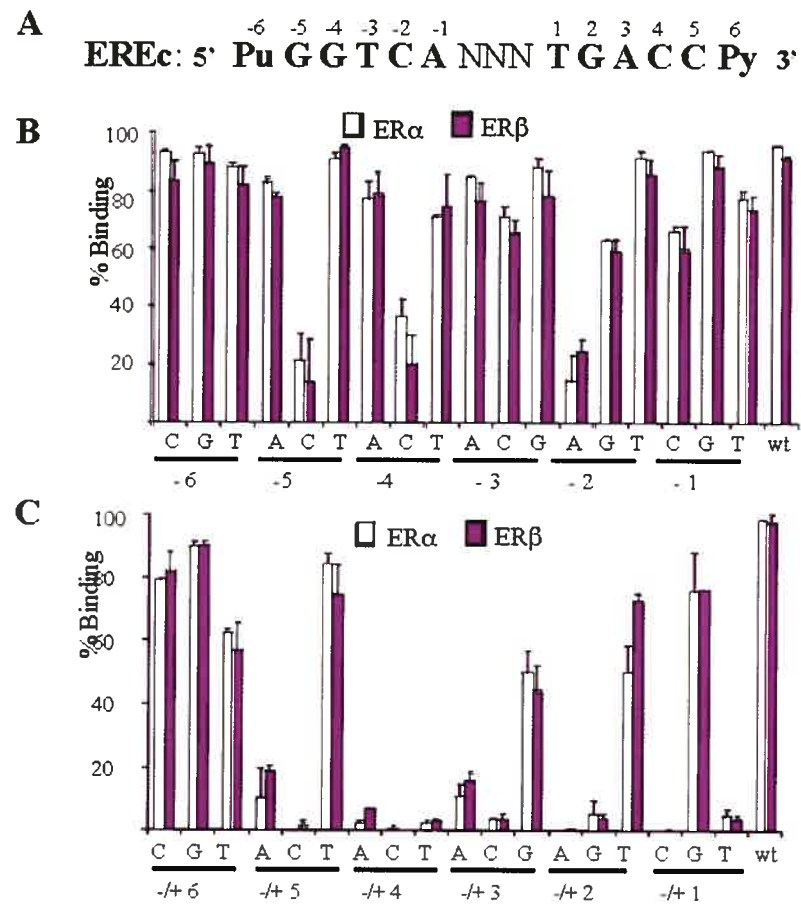


Figure 2

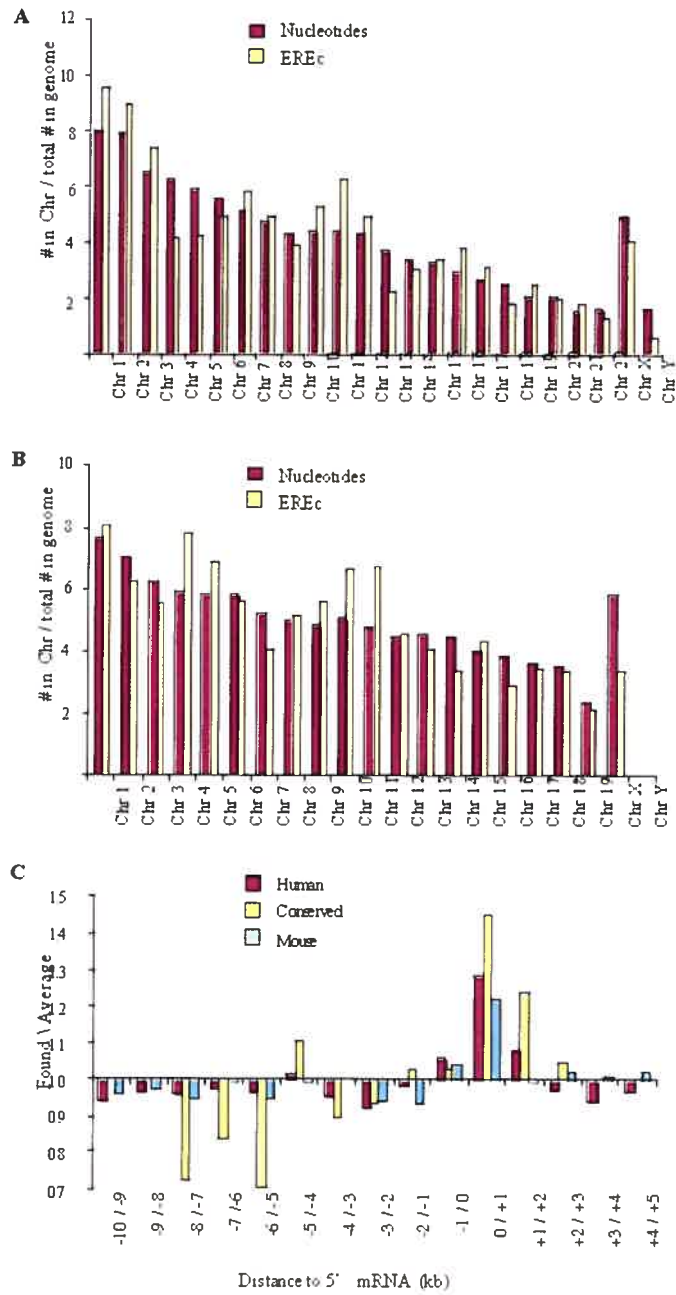


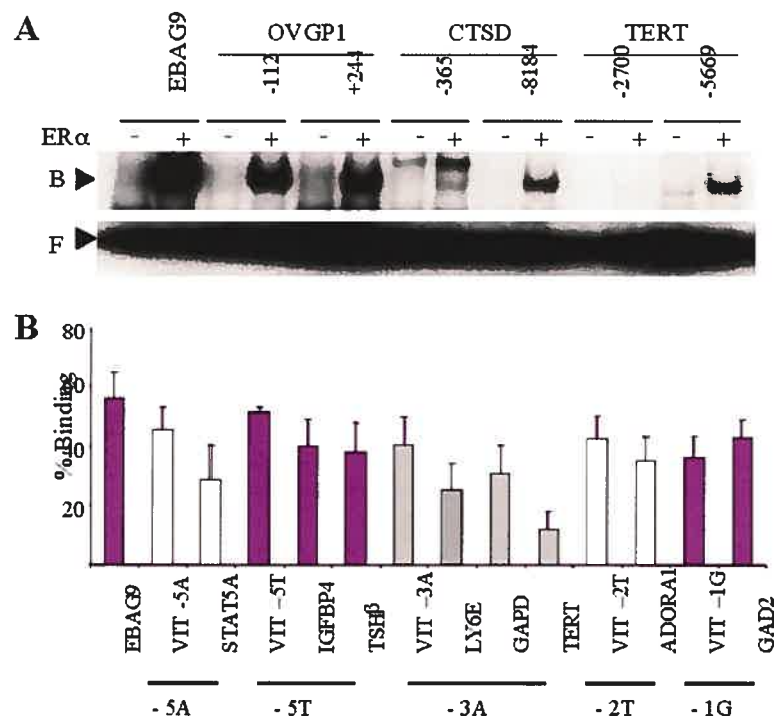
Figure 3

Figure 4

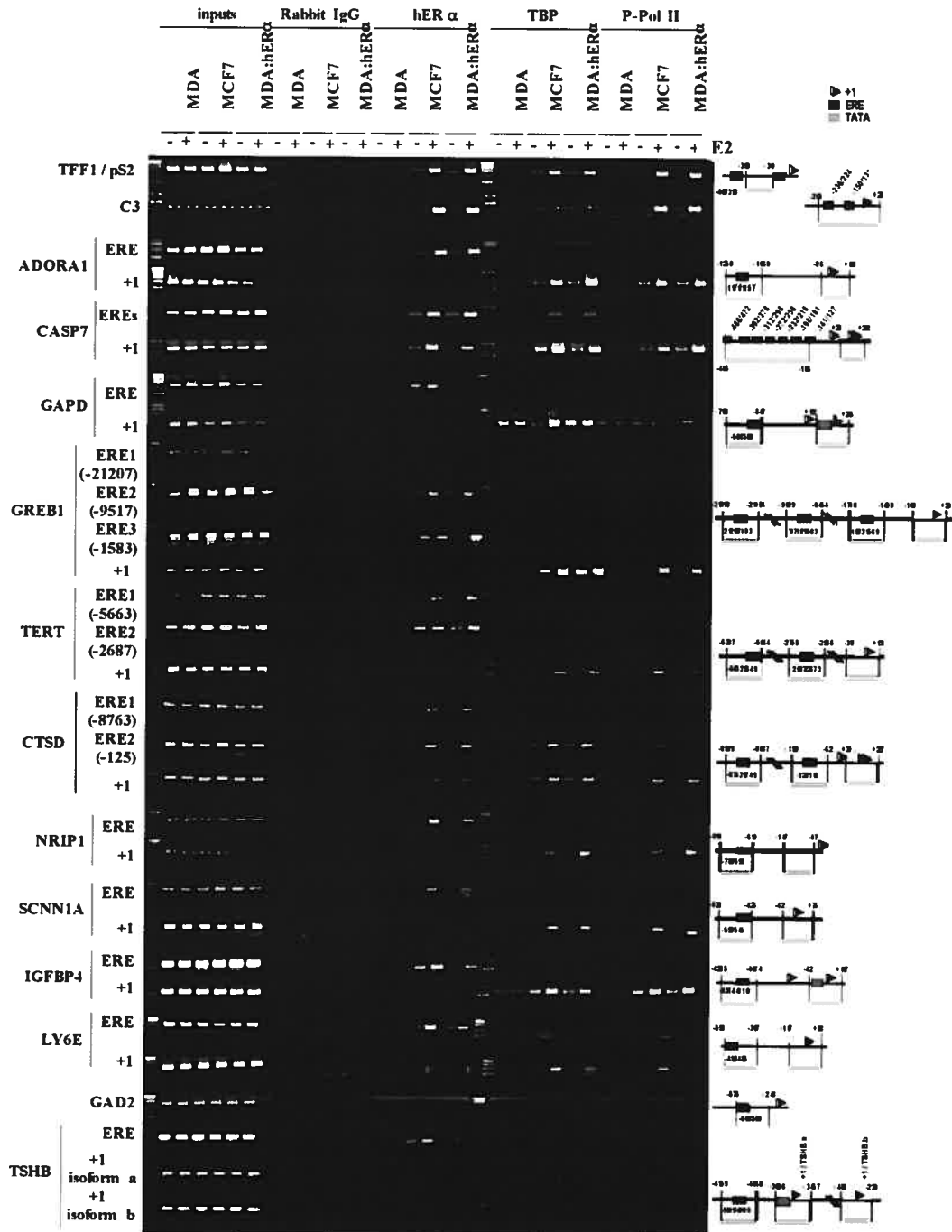


Table Supplémentaire 1

Human Gene Name	Chr#	Contig ID	Human ERE Sequence	Hs Distance			Mouse Gene Name	Chr#	Contig ID	Mouse ERE Sequence	Mm Distance			Difference		
				ERE-Genome	ERE-mRNA	ERE-CDS					ERE-Genome	ERE-mRNA	ERE-CDS	Min-Hs-ERE-Genome	Diff-mRNA	Diff-mHs-ERE-CDS
ABC3	Chr1	NT_021979	AGGTCAGGGTCACCC	4770	4732	4703	Abcd3	Chr3	NT_032940	ACGTCAAGGTCACCC	4320	4136	4136	450	596	573
AC02	Chr2	NT_011520	AGGCAAAAGGGTGTC	219	219	199	Aco2	Chr5	NT_036261	AGGTCAAGGCTGACCT	341	203	197	122	16	2
AC02	Chr1	NT_009327	CGGTCACTGGTCAACCT	344	-66	293	Asp2	Chr2	NT_032928	GGGTAAAGGGCTGACCT	-282	-282	-296	626	598	589
AC02	Chr1	NT_01295	GGGTCACTGGTCAACCT	5418	-369	293	Asp2	Chr2	NT_032928	GGGTAAAGGGCTGACCT	-282	-282	-296	626	598	589
ACTL5	Chr7	NT_007933	GGGCGACCACTGACCT	1012	991	918	Actl5	Chr9	NT_033942	GAAGTCAAGCTGACCA	-3709	-8798	-9323	1709	2099	2269
ACTL7A	Chr7	NT_003470	TGCTCAACTGTGACCC	3508	3508	3505	Actl7a	Chr9	NT_032961	AGGTGTCGGCTGACCT	-3044	-3044	-3044	464	464	474
ACTL7B	Chr7	NT_000473	GGGTCAAGCTGTGACCC	1036	1016	49	Actl7b	Chr9	NT_032961	AGGTGTCGGCTGACCT	1059	1059	972	27	27	923
ACV8R1B	Chr12	NT_029419	GGGTCAAGCTGTGACCT	-421	-436	-438	Actv8r1b	Chr15	NT_036261	TGCTCAAGGCTGACCT	-3176	-3206	-3215	1055	1060	1093
AD03	Chr1	NT_000470	GGGTCACTGTGTATGCT	13075	-4253	-10745	Ad03	Chr2	NT_032925	GGGTCAAGGCTGACCT	-5061	-5209	-16640	18136	956	5895
ADAMTS1	Chr1	NT_033899	TGGTCTCAAGCAGCACT	1198	117	19	Adamts1	Chr9	NT_033942	AGGTCAAGCTGTGCTCC	-1409	-1409	-1684	211	211	2221
ADAMTS2	Chr1	NT_033899	TGGTCTCAAGCAGCACT	1198	117	19	Adamts2	Chr9	NT_033942	AGGTCAAGCTGTGCTCC	-1409	-1409	-1684	211	211	2221
ADAT1	Chr16	NT_034797	AGGTCTTCTCTGACCA	-4181	-4101	-6558	Adat1	Chr8	NT_033967	GGTGTAAGTGTGACCT	-6310	-8211	-8239	2209	161	141
ADCY3	Chr2	NT_022184	GGGCGACAGGTGACCC	1482	777	630	Adcy3	Chr12	NT_035939	GGGTCAAGCTGTGCCC	1537	761	616	55	122	168
ADCY3	Chr2	NT_022184	GGGCGACAGGTGACCC	1482	777	630	Adcy3	Chr12	NT_035939	GGGTCAAGCTGTGCCC	1537	761	616	55	122	168
ADCY3	Chr2	NT_022184	GGGCGACAGGTGACCC	1482	777	630	Adcy3	Chr12	NT_035939	GGGTCAAGCTGTGCCC	1537	761	616	55	122	168
ADCY4	Chr7	NT_076437	AGGTCAAGCATGACCA	3652	3204	3288	Adcy4	Chr14	NT_036968	AGGCTCTACTGACCT	3265	3212	2871	477	82	417
ADCY5	Chr16	NT_010552	TGGTCTCACTGACCT	3261	3978	384	Adcy5	Chr16	NT_036964	ACCTGTCAGTGTGACCT	-2832	-2832	-2877	1216	109	1079
ADH7	Chr4	NT_016354	AGGTCATGTGTGACTT	4551	-4351	-4451	Adh7	Chr7	NT_033942	GGGTCAAGCTGTGACCA	-3819	-3933	-3948	532	412	503
ADORA3	Chr1	NT_019793	GGTTCAGCTGTGACCA	887	887	120	Adora3	Chr3	NT_032929	GGGTCAAGCTGACAG	711	711	334	176	176	214
ADORA3	Chr1	NT_019723	GGTTCAGCTGTGACCA	887	887	120	Adora3	Chr3	NT_032929	GGGTCAAGCTGACAG	486	456	109	401	401	11
ADORA3	Chr1	NT_019723	GGTTCAGCTGTGACCA	887	887	120	Adora3	Chr3	NT_032929	GGGTCAAGCTGACAG	486	456	109	401	401	11
ADORA3	Chr1	NT_019723	GGTTCAGCTGTGACCA	887	887	120	Adora3	Chr3	NT_032929	GGGTCAAGCTGACAG	486	456	109	401	401	11
AGPAT3	Chr2	NT_015115	TGGTCAAGCTGTGACCT	-1192	-1112	345	Agpat3	Chr10	NT_034946	AGGCGAGTGTGACCT	-1387	-8170	-13562	1509	1294	68472
AGRP	Chr16	NT_010498	TGGTCAAGCTGTGACCT	424	12	12	Agpr	Chr8	NT_034967	GGGTCAAGTATACCA	-548	-548	-548	972	360	557
AGRP	Chr16	NT_010498	TGGTCAAGCTGTGACCT	557	145	142	Agpr	Chr8	NT_034967	GGGTCAAGTATACCA	-548	-548	-548	972	360	557
AGRP	Chr16	NT_010498	TGGTCAAGCTGTGACCT	557	145	142	Agpr	Chr8	NT_034967	GGGTCAAGTATACCA	-548	-548	-548	972	360	557
AGT	Chr16	NT_010498	TGGTCAAGCTGTGACCT	557	557	142	Agpr	Chr8	NT_034967	GGGTCAAGTATACCA	-548	-548	-548	972	360	557
AGT	Chr16	NT_010498	TGGTCAAGCTGTGACCT	557	557	142	Agpr	Chr8	NT_034967	GGGTCAAGTATACCA	-548	-548	-548	972	360	557
AGT	Chr16	NT_010498	TGGTCAAGCTGTGACCT	557	557	142	Agpr	Chr8	NT_034967	GGGTCAAGTATACCA	-548	-548	-548	972	360	557
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKAP10	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8	NT_034967	GGGTCAAGCTGTGACCT	5219	98	66	5169	108	3347
AKP	Chr1	NT_021973	GGGTCAAGCTGTGACCT	50	10	-3281	Akp	Chr8								

Table Supplémentaire 1-suite

BRF1	Chr14	NT_026437	TGGCCAGCTGACCC	62165	-5539	47205	Brf1	Chr12	NT_039553	GGGTCAATTCACCC	32526	-4742	-5578	29639	797	52783
BRF1	Chr14	NT_026437	TGGCCAGCTGACCC	62165	-4993	47205	Brf1	Chr12	NT_039553	GGGTCAATTCACCC	32526	-4742	-5578	29639	251	52783
BRF2	Chr8	NT_007995	AGGTCAAGCTGGCCT	-7201	-7201	-7311	Brf2	Chr8	NT_039456	TGGTCAAGTACCC	-8246	-8246	-8252	10945	1045	1041
BSG	Chr19	NT_011255	TGTCAAATTTGATCC	-1377	-2627	-2684	Bsg	Chr10	NT_039496	GGGTCAAGTACCT	-2518	-2518	-2678	1141	109	6
C14orf11	Chr1	NT_026437	AGGACATACTGACCT	-1928	-1928	-2011	181001101dRuk	Chr12	NT_039549	GGTTCATGGCGACCT	94	41	6	2022	1969	2017
C1QA	Chr1	NT_004618	GGGTCAATTTACCA	2683	1731	1731	C1qa	Chr4	NT_039266	GGGTCAAGTTGACAG	1720	1720	710	963	11	1011
C1QG	Chr1	NT_004618	GGGTCAATTTACCA	4295	-4295	-4686	C1qg	Chr4	NT_039266	GGGTCAAGTTGACAG	-4233	-4233	-4565	62	41	121
C6orf11	Chr6	NT_007593	GGGTCAAGCTGACCC	-4952	-7031	-7075	Bing4	Chr17	NT_039469	GGGTCAATTTGACCC	-8160	-8218	-8218	1208	1187	1143
C6orf11	Chr6	NT_007593	GGGTCTCTGTGACCC	-6405	-6484	-6528	Bing4	Chr17	NT_039469	GGGTCAAGAGACCC	-7624	-7652	-7652	1219	1198	1154
C6orf17	Chr6	NT_034874	AGATCAGCTGACCA	-9610	-9610	-9925	Pcg-pending	Chr17	NT_039550	ATGTCAAGTTGACCA	-9636	-9636	-10278	26	26	353
C6orf17	Chr6	NT_034874	AGATCAGCTGACCA	-9610	-9610	-9925	Pcg-pending	Chr17	NT_039663	ATGTCAAGTTGACCA	-9822	-9822	-10464	212	212	539
C6orf17	Chr6	NT_033951	CGGTCAAAATGGCCC	3864	3371	3371	D17H6S56E.3	Chr17	NT_039649	GGGTCAATTTGACCA	3099	3099	2606	765	272	765
CA3	Chr8	NT_071037	AGGTCAATAAGACCT	-6702	-6702	-7383	Car3	Chr3	NT_039226	AGGTCAATTTGACCT	-5603	-5611	-5663	1099	1091	1720
CA7	Chr16	NT_010495	GGGTCACTGGGACCT	-7860	-7969	-7969	Car7	Chr8	NT_039467	GGGTCAACCTGACCA	-3821	-9939	-10008	4038	1960	2039
CACNA1F	ChrX	NT_011568	TGGTCAACCTGACCG	-3179	-3179	-3241	Cacna1f	ChrX	NT_039698	TGGTCAATGACCT	-2645	-2645	-2657	534	534	554
CACNA1F	ChrX	NT_011568	TGGTCAACCTGACCG	-456	-456	-718	Cacna1f	ChrX	NT_039698	TGGTCAATGACCT	-2645	-2645	-2687	1989	1919	1969
CACNA1F	ChrX	NT_011568	TGGTCAACCTGACCG	-456	-456	-718	Cacna1f	ChrX	NT_039698	TGGTCAATGACCT	-2645	-2645	-2687	1989	1919	1969
CACNA1F	ChrX	NT_011568	TGGTCAACCTGACCG	-456	-456	-718	Cacna1f	ChrX	NT_039698	TGGTCAATGACCT	-2645	-2645	-2687	1989	1919	1969
CACNA1F	ChrX	NT_011568	TGGTCAACCTGACCG	-456	-456	-718	Cacna1f	ChrX	NT_039698	TGGTCAATGACCT	-2645	-2645	-2687	1989	1919	1969
CALR3	Chr1	NT_011295	GGGTCAACAGACCA	2501	2501	2327	Calm3	Chr1	NT_039477	GGGTCAATGACCT	1421	1421	964	1080	1080	1363
CAPON	Chr1	NT_004668	TGGTCAAGCTGACCT	3889	3889	3487	633048P19Ruk	Chr8	NT_039466	AGGTCAATGACCT	-2478	-2478	-2573	432	432	400
CAPON	Chr1	NT_004668	TGGTCAAGCTGACCT	3889	3889	3487	633048P19Ruk	Chr1	NT_039185	AGGTCAATGACCT	2273	2273	2228	1616	1616	1259
CAPON	Chr1	NT_004668	TGGTCAAGCTGACCT	3889	3889	3487	633048P19Ruk	Chr1	NT_039185	AGGTCAATGACCT	2273	2273	2228	1616	1616	1259
CARD14	Chr17	NT_024871	GGGTGGTGGTGAACCT	13934	4616	4003	Card14	Chr1	NT_039521	AGGTCAAGCTGACCT	10591	3851	1553	3043	765	2450
CARD14	Chr17	NT_024871	GGGTGGTGGTGAACCT	13934	4616	4003	Card14	Chr1	NT_039521	AGGTCAAGCTGACCT	10591	3851	1553	3043	765	2450
CARD14	Chr17	NT_024871	GGGTGGTGGTGAACCT	13934	4616	4003	Card14	Chr1	NT_039521	AGGTCAAGCTGACCT	10591	3851	1553	3043	765	2450
CARD14	Chr17	NT_024871	GGGTGGTGGTGAACCT	13934	4616	4003	Card14	Chr1	NT_039521	AGGTCAAGCTGACCT	10591	3851	1553	3043	765	2450
CARD14	Chr17	NT_024871	GGGTGGTGGTGAACCT	13934	4616	4003	Card14	Chr1	NT_039521	AGGTCAAGCTGACCT	10591	3851	1553	3043	765	2450
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGTCAACTTGACCA	1458	1458	-4285	1850	2586	6375
CASP7	Chr10	NT_030259	GGGTGGGGTGAACCT	-392	-1128	-13160	Casp7	Chr19	NT_039692	GTGT						

Table Supplémentaire 1-suite

CP5F2	Chr14	NT_029437	AGGTGAGATTTGACCC	749	749	-8251	Cpef2	Chr12	NT_039552	AGGTGATCTTGACCTG	1525	-6535	-12494	776	7284	4243
CP5F4	Chr7	NT_007933	GCGTCACTACGACCT	-7778	-7878	-7888	Cpef4	Chr5	NT_039318	AGGTAGAGTGAACCT	-8497	-8497	-8630	719	619	747
CP5F4	Chr7	NT_007933	GGGTCACTACGACCT	-7778	-7878	-7888	Cpef4	Chr5	NT_039318	AGGTCACTGTAACCA	-7248	-7248	-7381	530	630	502
CP1C	Chr19	NT_011109	GGGCGCTCTGACCC	-395	-395	-1498	Cpic-pending	Chr7	NT_039420	AGGTGAGTGAATCC	-795	-835	-1742	408	440	244
CRELD1	Chr3	NT_005927	TGGTCAGAGTGGCT	-3844	-4107	-4440	Creld1	Chr5	NT_039355	CGGTCAAGGTGACCT	-4623	-4623	-4626	779	516	456
CRF	Chr17	NT_010748	AGGTTGAGGTGACCC	-1513	-1733	-1746	Clrf-pending	Chr11	NT_039521	GGGTGAGTGAACAT	-1568	-1568	-1568	55	165	178
CRHR1	Chr17	NT_010748	AGGTGAGGTGACCC	-7590	-7603	-7828	Crhr1	Chr11	NT_039521	AGGTCAATGTGACCC	-6786	-6959	-6989	804	644	839
CSH1	Chr17	NT_010783	TGGTCAGGTGACAT	-8972	-8972	-9083	Csh1	Chr13	NT_039578	GGTCAAGTGAACCT	-8046	-8046	-8109	926	926	979
CSMD1	Chr8	NT_023736	ATGTCAGATGACCT	-4055	-4160	-4445	Csmd1	Chr8	NT_039455	TGGTCTCTGACCC	-3898	-3898	-3898	157	262	547
CSNK1D	Chr17	NT_010663	GGGTGAGGTGATCT	-3142	-3142	-3458	Csnk1d	Chr11	NT_039521	GGTCAAGGTGACCC	-1878	-1889	-2198	1264	1253	1260
CSNK1E	Chr22	NT_011520	TGGTCAGGTGATCT	-75747	-5396	-8618	Csnk1e	Chr15	NT_039621	TGCTCAGAGTGAACCT	-5780	-5914	-9069	8127	518	451
CSNK1E	Chr22	NT_011520	TGGTCAGGTGATCT	-75747	-4691	-8618	Csnk1e	Chr15	NT_039621	TGCTCAGAGTGAACCT	-5780	-5914	-9069	8127	1223	451
CTSD	Chr11	NT_008310	GGGCAACATGACCC	-8182	-8763	-8818	Ctsd	Chr17	NT_039437	AGGCCAATGTAACCT	-1920	-8157	-8223	193	606	673
CUTL1	Chr7	NT_007933	TGGTCACTGACCT	1931	1879	1860	Cutl1	Chr11	NT_039314	AGGTCAACCTGACCT	197340	1783	-8245	195409	96	10105
CXCL16	Chr17	NT_010718	GGGTCTCCCTGACCC	-3807	-3905	-4328	Cxcl16	Chr11	NT_039520	AGGTCAAGGGGACCT	-4870	-5407	-5407	1063	1502	1079
CXCL2	Chr4	NT_006216	AGGTCTTTATGACCT	2704	2628	2553	Cxcl2	Chr11	NT_039307	TGGTCAAGGTGACCC	903	877	831	1801	1751	1722
CYBA	Chr16	NT_010542	AGGTCAAGGTGACCC	2768	2677	2649	Cyba	Chr8	NT_039467	GGGTGAGTGAACCT	2522	2508	2440	266	177	209
CYP1A1	Chr15	NT_010194	GGGTGAGGTGACCT	4242	4242	1803	Cyp1a1	Chr9	NT_039474	TGGTCACTTTGACCC	12216	3032	554	7974	1210	1249
CYP1A1	Chr15	NT_010194	GGGTCACTGTCACCT	4932	4932	2543	Cyp1a1	Chr9	NT_039474	TGGTCACTTTGACCC	12216	3032	554	7974	1950	1849
CYP1B1	Chr2	NT_022184	AGGCGCTTTGACCC	313	-105	-866	Cyp1b1	Chr17	NT_039658	CGGACAGATGACCT	425	367	-363	113	472	503
CYP1B1	Chr2	NT_022184	AGGCGCGGCGACCT	369	-49	-810	Cyp1b1	Chr17	NT_039658	CGGACAGATGACCT	425	367	-363	56	416	447
DIS155E	Chr1	NT_019273	GGGTAAATATGACCA	-718	-737	-1832	DJ181	Chr7	NT_039329	GGGTGAGTGAACCT	122	-8161	-18032	840	7424	779
DMS24E	Chr4	NT_006051	GGGTGCTCTGTCCT	38386	-931	-931	Nagl	Chr5	NT_039303	AGGTGAGCTGACCT	-663	-1565	-1565	39049	634	634
DMS24E	Chr4	NT_006051	GGGTGCTCTGTCCT	38386	-931	-931	Nagl	Chr5	NT_039303	GGGTGAGCTGACCT	853	-49	-49	37533	882	882
DAB2	Chr5	NT_006576	AGGTCAAGGTGACCA	4140	3755	-26769	Dab2	Chr15	NT_039617	TGGTCAAGATGACCT	91555	3863	-36481	87415	108	288
DCX	ChrX	NT_025319	ATGTCAGATGACCT	308	-178	-178	Dcx	Chr15	NT_039718	GGGTCAAGTGAACCT	607	390	-1400	209	1020	22
DCX	ChrX	NT_025319	ATGTCAGATGACCT	3062	2034	1286	Dcx	Chr15	NT_039718	GGGTCAAGTGAACCT	607	390	-1400	2455	1644	2668
DDA3	Chr1	NT_019273	AGGTCTTTATGACCT	-3835	-3835	-4284	Dda3-pending	Chr3	NT_039239	AGGTGAGCTGGCCT	-4911	-4950	-5374	1076	1115	1090
DDA3	Chr1	NT_019273	GGGTGAGCTGAGCT	-3005	-3005	-3454	Dda3-pending	Chr3	NT_039239	AGGTGAGCTGGCCT	-4911	-4950	-5374	1906	1945	1920
DDT	Chr22	NT_011520	GGGTGAGGTGACCC	47	47	-5369	Ddt	Chr10	NT_039510	TAGTCACTCTGACCC	-4984	-5005	-5056	5031	5052	313
DDT	Chr22	NT_011520	GGGTGAGGTGACCC	47	47	-5369	Ddt	Chr10	NT_039496	GGGTGAGATGACCT	-4971	-4992	-5043	5018	5039	326
DEF6	Chr6	NT_007592	AGGTCAAGCTGTCCT	-4069	-4108	-4108	Def6	Chr17	NT_039469	AGGTCAAGCTGTCCT	-4721	-4746	-4759	652	638	651
DEF6	Chr6	NT_007592	GGGTCTCTGTCACCT	-2960	-2939	-2939	Def6	Chr17	NT_039469	AGGTCAAGCTGTCCT	-4721	-4746	-4759	1821	1807	1820
DEF6	Chr6	NT_007592	GGGTCTCTGTCACCT	-2960	-2939	-2939	Def6	Chr17	NT_039469	AGGTCAAGCTGTCCT	-4721	-4746	-4759	563	591	594
DES	Chr2	NT_005403	AGGTGAGATGACCT	1894	1888	1808	Dcs	Chr1	NT_039171	AGGTGATGTAACCT	1929	1919	1915	95	101	107
DF	Chr19	NT_011255	AGGTCAAGGTGACCC	3590	3543	2617	Adn	Chr10	NT_039496	GGGTCACTGTCGACCA	3515	3481	3467	75	62	850
DGCR8	Chr22	NT_011519	AGGTCAAGGTGACCC	-8565	-8565	-14218	Dgcr8	Chr16	NT_039624	AGGTGAGCTGATCT	-6830	-6830	-11660	1735	1735	2558
DGCR8	Chr22	NT_011519	AGGTCAAGGTGACCC	-8565	-8565	-14218	Dgcr8	Chr16	NT_039624	AGGTGAGCTGATCT	-6830	-6830	-11660	1737	1737	2558
DHCHD	Chr16	NT_010498	AGGTCACTGTCGACCA	-2689	-2689	-5979	Dhchd	Chr8	NT_039467	AGGTCAAGGTGACCC	-3158	-3453	-3459	469	764	2520
DHCHD	Chr16	NT_010498	AGGTCACTGTCGACCA	-2689	-2689	-5979	Dhchd	Chr8	NT_039467	AGGTCAAGGTGACCC	-3158	-3453	-3459	3802	4097	813
DIRA1	Chr19	NT_011255	AGGTCAAGGTGACCC	-4791	-4791	-10351	Dira1	Chr10	NT_039496	AGGTCACTGTCGACCA	-5302	-5302	-8262	1489	1489	2039
DIRA1	Chr19	NT_011255	AGGTCAAGGTGACCC	-4791	-4791	-10351	Dira1	Chr10	NT_039496	AGGTCACTGTCGACCA	-5302	-5302	-8262	1155	1155	555
DKFZP566C23	Chr6	NT_006713	TGGTCACTGTCGACCT	4471	-4491	-7646	D17Wm94e	Chr17	NT_039655	CGGTCAAGTGAACCT	3764	3764	3615	998	1028	1028
DKFZP761J13	Chr5	NT_006713	TGGTCACTGTCGACCT	4471	-4491	-7646	D17Wm94e	Chr17	NT_039655	CGGTCAAGTGAACCT	3764	3764	3615	1707	156	2972
DLG3	ChrX	NT_011669	AGGTCAAGGTGACCC	3878	3791	3558	Dlgh3	Chr1	NT_039711	GGGTCAAGCTGTCCT	3717	3717	3737	161	74	161
DLX2	Chr2	NT_005403	AGGTCACTGTCGACCT	-7943	-8151	-8151	Dlx2	Chr2	NT_039208	AGGTCAAGGTGACCT	-9415	-9415	-9764	1472	1264	1613
DLX4	Chr17	NT_010783	TGGTCAAGTGAACCT	261	261	-8	Dlx4	Chr11	NT_039521	GGGTCAAGTGAACCT	114405	298	298	114444	37	306
DMR5	Chr19	NT_011109	AGGTCAAGGTGACCC	-6631	-7115	-7893	Dmr5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	1218	1534	756
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	2070	1242	1488
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	111	78	8913
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	3501	60311	
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	393	220	1109
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	2243	2421	1532
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	1954	1773	1807
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	479	474	474
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	127	127	127
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	714	23703	
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	1719	1719	1719
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	327	349	349
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	1583	1454	1603
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	623	623	623
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	593	4068	6640
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	226	4582	5816
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	2987	2431	1837
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	1794	1794	1794
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	1834	1898	1858
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	518	175	16676
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	518	217	16676
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	38	156	0
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	912	822	885
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	290	27	144
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	488	231	317
DNAJB5	Chr7	NT_008413	ATGTCAGATGACCT	-8452	-8532	-11978	Dnajb5	Chr7	NT_039400	AGGTCAAGGTGACCT	-8249	-8649	-8649	524	640	623
DNAJB5	Chr7</															

Table Supplémentaire 1-suite

FBXL12	Chr19	NT_011295	GTGTACAGTTGACCA	-4985	-5009	-5251	Pfbx12	Chr9	NT_039472	GGGTCAAAATGCCCC	-2756	-6857	-7256	2229	1848	2005
FBXL12	Chr19	NT_011295	AGCTACTGTGACCT	-4362	-4386	-4628	Pfbx12	Chr9	NT_039472	GGGTCAAAATGCCCC	-2756	-6857	-7256	1606	2471	2628
PGF10	Chr5	NT_006576	TGTTCCCGCTGACCT	1095	303	303	Fg10	Chr13	NT_039590	AGGTCAAGCGGACCA	1013	292	292	82	11	11
PGF2	Chr4	NT_016354	AGGTTACAAGTACGCC	-3472	-3472	-3774	Fg2	Chr13	NT_039228	AGGTCACTTTGACCA	-4951	-5126	-5126	1479	1654	1352
PGF2	Chr4	NT_016354	AGGTTACAAGTACGCC	-3472	-3472	-3774	Fg2	Chr13	NT_039248	AGGTCACTTTGACCA	-5174	-5349	-5349	1702	1877	1575
PGF2	Chr4	NT_016354	AGGTGGGCGCTGACCC	4284	4284	3982	Fg2	Chr13	NT_039228	GGGTCACTTTGACCC	4185	4010	4010	99	274	28
PGF2	Chr4	NT_016354	AGGTGGGCGCTGACCC	4284	4284	3982	Fg2	Chr13	NT_039248	GGGTCACTTTGACCC	4153	3978	3978	131	306	4
PGF8	Chr10	NT_030059	GGGTCAAGCTGCTGCC	67	67	-103	Fg8	Chr13	NT_039606	TGTTCACTTTGACCT	2134	-237	-375	2338	33	6
PGF8	Chr10	NT_030059	GGGTCAAGCTGCTGCC	2392	-204	-381	Fg8	Chr19	NT_039695	GGGACGGCGTGACCC	191	149	-50	124	82	53
PGF9	Chr13	NT_009799	TGTTCACTTTGACCT	-204	-204	-103	Fg9	Chr19	NT_039692	GGGTCAAGCTGACCC	195	153	-36	128	86	67
FBP	Chr11	NT_033903	AGGTCACTTTGACCA	1358	1314	1303	Fbp	Chr14	NT_039684	GGGTCAAGCTGACCC	1085	1007	1007	178	254	118
FKBP4	Chr12	NT_009759	AGGGCATCTTGACCT	1099	1065	912	Fkbp4	Chr6	NT_039356	GGGTCAAGGCGGCTCC	921	811	794	307	307	296
FLJ10055	Chr17	NT_010641	AGGTCAAGTGAACCT	-6287	-6315	-6351	D11Etd498e	Chr11	NT_039521	GAGTCAGCTGACCT	-5502	-5559	-5628	779	756	723
FLJ10477	Chr8	NT_008251	AGGTGCTGTGATCC	-7068	-7068	-7299	483341A01Ruk	Chr8	NT_039456	GGGTCACTATGACCA	-5600	-5600	-5829	1468	1468	1470
FLJ10709	Chr1	NT_077914	AGGGCAGCTGACCA	-5737	-5777	-5843	2400004H09Ruk	Chr4	NT_039268	AGGTCAAGTGAACCT	-6908	-6908	-6989	1171	1131	1155
FLJ12875	Chr1	NT_006100	AGTTGCACTGACCT	-8563	-8563	-8700	0610009K11Ruk	Chr4	NT_039266	AGGTCACTTTGACCT	-9054	-9054	-9123	491	491	423
FLJ14399	Chr11	NT_033899	AGGTCAAGCTGACCA	2392	2385	1782	061003901Ruk	Chr4	NT_039473	AGGTCACTTTGACCT	3569	3478	2995	1177	1093	1213
FLJ14981	Chr19	NT_011255	GGGTGAATTTGACCA	2829	2829	2789	E130379M08Ruk	Chr17	NT_039456	AGGTCAAGCTGACCT	4813	4718	4562	1904	1889	1773
FLJ22170	Chr17	NT_010718	TGGGCAAGCTGACCT	5842	-7768	-9845	1500010D02Ruk	Chr11	NT_039520	AGGTCAAGCTGACCT	-8110	-8110	-16452	13952	342	6907
FLJ22457	Chr1	NT_019273	AGGTCAAGTGAACCA	7153	3277	3077	2010303M01Ruk	Chr3	NT_039239	TGTTCAAGCTGACCT	5585	1550	1346	1568	1777	1731
FLJ25082	Chr10	NT_008583	AGGTCAAGCTGCTCC	-1380	-1380	-40200	1700010P07Ruk	Chr14	NT_039596	GGGTGAGCTGACCC	399	399	323	1779	1779	40523
FLJ31121	Chr5	NT_029289	TGTTCAAGTTGACCC	-4625	-4695	-4709	2610510D14Ruk	Chr18	NT_039674	TGTTGAGTTTGAACCC	-6011	-6015	-6029	1386	1370	1320
FLJ31523	Chr17	NT_024871	AGGTCAAGCTGACCA	4114	4114	4092	241002010Ruk	Chr11	NT_039521	TGTTCACTTTGAACCC	3529	3529	3477	585	585	615
FLJ22771	Chr11	NT_032983	CTGTCACTGACCT	1088	1088	982	4930579J09Ruk	Chr19	NT_039687	TGTTCAAGTTGACCT	783	752	196	305	336	202
FMNL	Chr17	NT_010748	AGGTCAAGCTGACCT	8500	-2943	-3234	Fmnl	Chr11	NT_039521	TGTTGAGGAACTGACCT	-1357	-1620	-1620	9857	423	1614
FN3K	Chr17	NT_010663	GGGTCAAGTGAACCT	-2832	-2832	-2834	Fn3k	Chr11	NT_039521	TGTTGAGGAACTGACCT	-1452	-1473	-1473	1335	1380	1361
FOXC1	Chr6	NT_034880	GGGTCTTTGACCC	-6907	-6907	-6907	Foxc1	Chr13	NT_039578	TGTTCTTTGACCT	-6419	-6419	-6629	488	488	52
FZD6	Chr8	NT_008046	GGGTCAATATGACCA	1367	1355	137	Fzd6	Chr15	NT_039621	GGGTCAATATGACCA	1191	270	137	176	1085	0
GFPD	ChrX	NT_025965	AGGTCACTGCTGACCT	4556	4556	3457	Gfpdx	ChrX	NT_039740	AGGTCACTATGACAC	3268	3565	2870	688	991	587
GFPD	ChrX	NT_025965	AGGTCACTGCTGACCT	4556	4556	3457	Gfpdx	ChrX	NT_039708	AGGTCACTATGACAC	3798	3495	2800	758	1061	657
GFPD	ChrX	NT_025965	AGGTCACTGCTGACCT	4556	4556	3457	Gfpdx	ChrX	NT_039740	AGGTCACTATGACAC	3273	2970	2275	1283	1586	1182
GAC1	Chr1	NT_034410	TGGTAAACTGACCC	-7796	-7912	-7323	Lrrn2	ChrX	NT_039708	AGGTCACTATGACCT	3203	2900	2205	1353	1656	1252
GAC1	Chr1	NT_034410	TGGTAAACTGACCC	-7740	-7856	-7327	Lrrn2	Chr1	NT_039180	GGGTCACTATGACCT	-7482	-7482	-64157	314	300	9116
GAD2	Chr10	NT_008705	AGGTCAAGCTGACCT	-339	-564	-709	Gad2	Chr1	NT_039356	AGGTCACTATGACCT	-7482	-7482	-64157	314	300	9116
GAD2	Chr10	NT_008705	AGGTCAAGCTGACCT	-339	-564	-709	Gad2	Chr2	NT_039205	AGGTCAAGCTGACCT	-283	-349	-694	56	215	15
GALE	Chr1	NT_004391	GGGTCACTATGACCT	-7213	-8558	-8987	Gale	Chr2	NT_039205	GGGTCAAGCTGACCT	-1428	-1494	-1839	1089	930	1130
GALNT9	Chr12	NT_035243	AGGTCAAGCTGACCT	-3179	-3415	-5517	Galn9	Chr4	NT_039266	AGGTCAATTTTACCA	-5507	-6042	-7116	1706	1616	1871
GALNT9	Chr12	NT_035243	AGGTCAAGCTGACCT	-3179	-3415	-5517	Galn9	Chr11	NT_039520	TGTTCAAGTATGCC	118069	-5978	-5999	121248	2563	482
GALR2	Chr17	NT_010641	GGGTCACTTTGACCT	4732	4496	2394	Galn9	Chr11	NT_039520	GGGTGAGCTATGACCC	126063	2014	1995	121331	2400	399
GAPD	Chr12	NT_009759	AGGTCAAGCTGACCT	-720	-738	-1053	Gapd	Chr11	NT_039521	TGTTCAAGCTGACCT	2203	1681	1681	2062	1515	1588
GAS7	Chr17	NT_010718	AGGGCAGCAGGACCC	237823	-1169	-1312	Gas7	Chr11	NT_039520	AGGTCACTTTGACCT	-1067	-1067	-1171	238900	102	74
GAS7	Chr17	NT_010718	AGGGCAGCAGGACCC	237823	-1169	-1312	Gas7	Chr11	NT_039520	AGGTCACTTTGACCT	-1067	-1067	-1171	240101	1109	1070
GATA6	Chr18	NT_010966	GGGCGCTGCTGACCC	-433	-2231	-2573	Gata6	Chr18	NT_039674	GGGCGCTGCTGACCC	-2456	-2456	-2472	2023	225	101
GBA2	Chr9	NT_008413	GGGTCCCATGACCT	533	533	9	Gba2	Chr4	NT_039260	GGGTCCCATGACCC	495	187	-15	38	346	24
GBA2	Chr9	NT_008413	GGGTCCCATGACCT	533	533	9	Gba2	Chr4	NT_039260	GGGTCCCATGACCC	495	187	-15	1123	1431	1109
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCC	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1618	1094	Gccr	Chr11	NT_039521	GGGTCCCATGACCT	6000	2666	2784	4661	727	820
GCCR	Chr17	NT_010718	GGGTCCCATGACCT	1618	1											

Table Supplémentaire 1-suite

HISTH1A	Chr6	NT_007592	AGGTACGGGCGTCCC	-9248	-9248	-9304	HistH1a	Chr13	NT_039578	TGGTCAATGATGTC	-9702	-9702	-9751	454	454	447
HKEA	Chr6	NT_007592	GGGTCAATGACGACCT	504	412	86	H2K4	Chr17	NT_039469	AGGTACCAATGACCC	648	400	120	144	144	1400
HLXB9	Chr7	NT_007741	TGGTCAAGTTCACCT	-6098	-6334	-6334	Hlxb9	Chr5	NT_039300	GGGTACAGGTGACGG	-7580	-7429	-7424	1482	1482	1490
HMGCL	Chr1	NT_004391	AGGTCTCATGACCC	297	297	290	Hmgcl	Chr4	NT_039266	GGGTCAAGCTGACCT	-1185	-1185	-1221	1482	1482	1511
HNRPAB	Chr5	NT_077451	GAGTCAGCGTGACCC	2194	2161	1863	Hnrapab	Chr11	NT_039520	AGGTCTTACTGACCT	2538	2527	2254	344	366	391
HNRPDL	Chr4	NT_016354	TGGTCAGGCGTACAC	4639	4639	4059	Hnrapdl	Chr5	NT_039308	TGGTCAGGCGTACAC	4121	3226	3102	518	1413	957
HNRPDL	Chr9	NT_011109	TGGCCAGCGTGACCT	-6342	-6779	-8807	Hnrapl	Chr7	NT_039413	AGGTCAAAATGACAC	-7049	-9174	-9265	727	395	458
HOXA9	Chr7	NT_007819	AGGTCAATGCGCACCC	3048	-1900	-1973	Hoxa9	Chr6	NT_039343	GGGTCAAGTTGCCCG	2248	2248	982	820	4148	2955
HOXB1	Chr7	NT_010783	CTGTCAAGGTGACCT	4301	4301	4295	Hoxb1	Chr11	NT_039521	GGGTCAAGTTGACAG	3814	3814	3739	487	487	556
HOXB5	Chr17	NT_010783	AGGCCAGCGTGACCC	-5854	-6059	-6133	Hoxb5	Chr11	NT_039521	AGGCATTTGTGACCA	-5527	-5527	-6546	327	523	413
HOXB5	Chr17	NT_010783	AGGCCAGCGTGACCC	-5854	-6059	-6133	Hoxb5	Chr11	NT_039521	GGGTACGCTGGGCT	-5116	-5116	-6135	738	943	2
HOXB6	Chr17	NT_010783	AGGCCAGCGTGACCC	5157	-1361	-1665	Hoxb6	Chr11	NT_039521	GGGTACGCTGGGCT	-1476	-1692	-1698	6633	331	33
HOXB6	Chr17	NT_010783	AGGCCAGCGTGACCC	5157	-1361	-1665	Hoxb6	Chr11	NT_039521	AGGCATTTGTGACCA	-1887	-2103	-2109	7044	742	444
HOXC12	Chr2	NT_029419	AGGTCTTTCTGACCT	-5201	-5201	-5201	Hoxc12	Chr12	NT_039621	AGGTCTTTCTGACCT	-5113	-5113	-5113	88	88	88
HOXD13	Chr2	NT_029403	TGGTCAAAATGACCC	-8865	-8865	-8976	Hoxd13	Chr12	NT_039208	TGGTCAAAATGACCC	-8719	-8719	-8719	146	146	257
HOXD14	Chr2	NT_029403	TGGTCAAAATGACCC	2556	1254	-1505	Hoxd14	Chr12	NT_039208	GGGTCAAAATGACCA	13134	3074	-2225	14320	4332	720
HSCARG	Chr6	NT_010552	GGTCTCCAGTGACCT	2307	904	141	H10025F24Ruk	Chr16	NT_039424	AGGTCAAAATGACCA	698	645	520	1609	259	379
HS11B2	Chr16	NT_010498	GGTCTCCAGTGACCA	-7942	-8412	-8522	Hsd11b2	Chr8	NT_039467	AGGTCAATCTGACCT	-8591	-8658	-8669	649	274	177
HS17B8	Chr6	NT_007592	GGGTCAATGACGACCT	-3310	-3310	-3337	H2K4	Chr17	NT_039469	AGGTCAATGACGACCT	-3101	-3101	-3111	209	209	226
HSPA1B	Chr6	NT_007592	AGATCAAGGTGACCT	1621	1621	1417	Hspa1b	Chr17	NT_039469	AGGTCAATGACGACCT	509	509	-1408	1112	1112	2823
HSPA1B	Chr6	NT_007592	AGATCAAGGTGACCT	-2109	-2144	-5201	Hspa1b	Chr17	NT_039469	AGGTCAATGACGACCT	-2209	-2238	-6493	100	94	1292
HSPA8	Chr1	NT_033899	GGGTCAATGACGACCT	-2702	-2702	-3514	Hspa8	Chr9	NT_039473	GGGTCAATGACGACCT	-2380	-2345	-2345	322	543	260
HSPA8	Chr1	NT_033899	GGGTCAATGACGACCT	-2702	-2702	-3514	Hspa8	Chr9	NT_039473	AGGTCAATGACGACCT	-1740	-1741	-1741	962	1773	1773
HSPB7	Chr1	NT_077957	TGGTCTATGTCGCC	-1971	-3109	-3602	Hspb7	Chr4	NT_039266	TGGTCTATGTCGCC	-1377	-4327	-4383	1406	1218	781
HSPC111	Chr5	NT_023133	TGGTCAAGGTGACCT	-6966	-6966	-10509	D13Wsu177e	Chr13	NT_039586	GGGTCAAGGTGACCT	-9186	-9186	-9289	2469	2469	1759
HSPC111	Chr5	NT_023133	AGGTCAAGGTGACCT	-6717	-6717	-7530	D13Wsu177e	Chr13	NT_039586	GGGTCAAGGTGACCT	-9186	-9186	-9289	2469	2469	1759
HTR3B	Chr1	NT_033899	TGGTCAATGACGACCT	3388	3324	3257	Htr3b	Chr9	NT_039473	TGGTCAAGGTGACCT	4655	4487	4450	1267	1163	1193
HTR3B	Chr1	NT_033899	TGGTCAATGACGACCT	3388	3324	3257	Htr3b	Chr9	NT_039473	AGGTCAATGACGACCT	4997	4829	4792	1609	1505	1535
ICAM4	Chr19	NT_011295	TGGCCAAAGTGACCT	1499	-5149	-9305	Icam4	Chr9	NT_039473	AGGTCAATGACGACCT	657	657	657	80	43	34
ICAM4	Chr19	NT_011295	TGGCCAAAGTGACCT	737	700	691	Icam4	Chr9	NT_039473	AGGTCAATGACGACCT	1754	1754	1754	1054	1054	1054
ICAM5	Chr19	NT_011295	TGGCCAAAGTGACCT	-2275	-2275	-2340	Icam5	Chr9	NT_039473	AGGTCAATGACGACCT	-1952	-1979	-2019	323	296	321
ICAM5	Chr19	NT_011295	TGGCCAAAGTGACCT	-2275	-2275	-2340	Icam5	Chr9	NT_039473	GGGTCAAGGTGACCT	-855	-882	-922	1420	1393	1418
IFNA4	Chr9	NT_008413	GGGTCAAGGTGACCT	-3609	-3609	-3749	Iffa4	Chr4	NT_039266	GGTTCACATGACCT	-5212	-5212	-5212	1603	1603	1463
IFRD1	Chr7	NT_007933	GGGTCAAGGTGACCT	30247	2912	2702	Ifrd1	Chr12	NT_039548	AGGTCAAGGTGACCT	2615	2597	2389	27632	315	313
IKBK1	Chr2	NT_029965	AGGTCAATGACGACCT	1499	-5149	-9305	Ikbk1	Chr2	NT_039740	AGGTCAATGACGACCT	1284	-6909	-6924	215	170	2381
IKBK1	Chr2	NT_029965	AGGTCAATGACGACCT	1499	-5149	-9305	Ikbk1	Chr2	NT_039740	AGGTCAATGACGACCT	689	-7504	-7519	810	2355	1766
IL10	Chr1	NT_021877	TGGTCAATGACGACCT	-336	-336	-395	Il10	Chr1	NT_039180	AGGTCAATGACGACCT	-355	-355	-420	382	376	376
IL10	Chr1	NT_021877	AGGTCAATGACGACCT	-336	-336	-395	Il10	Chr1	NT_039180	AGGTCAATGACGACCT	-355	-355	-420	382	376	376
IL17B	Chr5	NT_029289	GGGTCAAGGTGACCT	16005	-8922	-8972	Il17b	Chr18	NT_039674	GGGTCAAGGTGACCT	-5280	-8330	-8379	21285	592	592
IL18BP	Chr1	NT_033927	GGGTCAAGGTGACCT	2638	1543	615	Il18bp	Chr7	NT_039429	TGGTCAATGACGACCT	1382	847	6	1256	696	609
IL18BP	Chr1	NT_033927	GGGTCAAGGTGACCT	2638	1543	615	Il18bp	Chr7	NT_039429	TGGTCAATGACGACCT	1382	847	6	1256	696	609
IL1RAPL2	Chr2	NT_011651	AGGTCAAGGTGACCT	194	194	194	Il1rapl2	Chr2	NT_039716	AGGTCAATGACGACCT	-440	-440	-440	634	634	634
IL1RAPL2	Chr2	NT_011651	AGGTCAAGGTGACCT	194	194	194	Il1rapl2	Chr2	NT_039716	AGGTCAATGACGACCT	-440	-440	-440	634	634	634
ILK	Chr1	NT_023310	AGGTCAAGGTGACCT	2598	2570	2080	Ilk	Chr7	NT_039429	TGGTCAATGACGACCT	2565	2565	2164	5	31	84
ILK	Chr1	NT_023310	AGGTCAAGGTGACCT	2598	2570	2080	Ilk	Chr7	NT_039429	TGGTCAATGACGACCT	2565	2565	2164	5	31	84
INCEP	Chr1	NT_033903	AGGTCAAGGTGACCT	5194	1030	1030	Incep	Chr19	NT_039687	AGGTCAAGGTGACCT	752	670	-3339	4442	360	4369
IRF4	Chr6	NT_034880	AGGTCAAGGTGACCT	-6638	-6638	-8052	Irf4	Chr13	NT_039578	GGGTCAAGGTGACCT	-7143	-7143	-8465	505	484	413
IRS1	Chr2	NT_005403	AGGTCAAGGTGACCT	-462	-462	-1483	Irs1	Chr1	NT_039172	AGGTCAAGGTGACCT	-795	-795	-1519	333	333	36
JAG2	Chr14	NT_026437	GGGTCAAGGTGACCT	-1771	-1771	-2175	Jag2	Chr12	NT_039553	GGGTCAATGACCT	-3659	-3659	-3670	1888	1888	1495
JMI	Chr2	NT_011568	GGGTCAAGGTGACCT	1071	1071	915	D3Xim40e	Chr2	NT_039688	TGGTCAATGACCT	961	878	834	110	123	81
KAI1	Chr1	NT_009237	AGGTCAAGGTGACCT	1324	2366	-28739	Kai1	Chr2	NT_039740	AGGTCAATGACGACCT	2677	2176	-2275	353	90	3954
KARS	Chr6	NT_024797	GGTCAAGGTGACCT	550	550	510	Kars	Chr6	NT_039467	TGGTCAATGACCT	-568	-639	-639	1118	1118	1118
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT_039356	AGGTCAATGACGACCT	-2151	-2151	-3608	1766	1766	3223
KCNAB1	Chr12	NT_009759	AGGTCAAGGTGACCT	-385	-385	-385	Kcnab1	Chr6	NT							

Table Supplémentaire 1-suite

LTF	Chr3	NT_005825	AGGTGACCTGACCA	-1166	-1166	-1462	Ltf	Chr9	NT_039480	AGGTCAAGGTAACCC	-340	-340	-374	826	826	1088
LTF	Chr3	NT_005825	AGGTGACCTGACCT	81	81	-377	Ltf	Chr9	NT_039480	AGGTCAAGGTAACCC	-340	-340	-374	259	259	3
LUM	Chr12	NT_019546	AGGTGAGGTTTACCA	3168	3024	486	Lum	Chr10	NT_039480	AGGTCAAGGTAACCC	3277	2898	472	109	126	14
LY6G6C	Chr6	NT_007592	AGGTGCGGTGAGCC	4661	4549	405	Ly6g6c	Chr17	NT_039462	AGGTCAAGTGAACCA	6750	4749	4749	2029	209	254
MADZL1	Chr4	NT_016354	AGGTGAGCTTGACCC	-1255	-1521	-1595	Madz1	Chr6	NT_039434	AGGTCAAGGAGGACCT	-284	-373	-421	971	1148	1174
MADCAM1	Chr19	NT_010255	AGGTGAGGCTGACCA	4795	-398	-398	Madcam1	Chr10	NT_039496	AGGTCAAGGTAAGCC	-539	-539	-559	5334	141	161
MAFG	Chr17	NT_010663	GGGTGCTGATAAACCC	4543	4516	9	Mafg	Chr11	NT_039521	GGGTGCTGATAAACCC	3574	9	9	969	4507	0
MAF2A2	Chr15	NT_010274	GGGTCAATTTGACCT	4141	4141	2859	Man2a2	Chr7	NT_039428	AGGTCAACATGACCT	3602	3602	2200	539	539	659
MAP2B3	Chr17	NT_030843	GGGTGAGGCTGACCC	14301	629	629	Map2b3	Chr11	NT_039520	GGGTGAGGTTGTCCT	1505	1505	1339	12796	876	710
MAP2C7	Chr19	NT_007812	GGGTCAAGGCTGACCC	372	335	270	Map2c7	Chr6	NT_039455	GGGTCAAGGCTGACCT	435	189	-424	63	146	444
MAP3K5	Chr6	NT_025741	TGGTCAAGGCTGATCT	-7759	-7759	-8120	Map3k5	Chr10	NT_039491	GGGTCAAAATGACCT	-9372	-9372	-9372	1613	1613	1252
MAP4K2	Chr11	NT_033903	GGGTGATGCGGACCC	1082	1082	990	Map4k2	Chr19	NT_039604	AGGTCAAGGCTGACCT	2954	2864	2864	1872	1782	1474
MAPK11	Chr22	NT_019197	GGGTGCTGAGTACCC	-2527	-2527	-2612	Mapk11	Chr15	NT_039621	AGGTCAAGCTGATCT	-2400	-2442	-2515	127	85	97
MAPK11	Chr22	NT_019197	GGGTGCTGAGTACCC	-513	-513	-598	Mapk11	Chr15	NT_039621	AGGTCAAGCTGATCT	-2400	-2442	-2515	1887	1929	1917
MAFN2	Chr8	NT_020846	GGGTGATGATGACCA	10601	-8284	-8284	Mafrn2	Chr15	NT_039620	GGGTGAGGTTGACCT	445	445	-10116	10156	8729	1708
MAZ	Chr16	NT_024812	AGGTCAAGGCTGACCT	-7044	-7145	-7188	Maaz	Chr7	NT_039433	GGGTGAGCTGACCT	-7032	-7032	-7515	12	113	329
MB	Chr22	NT_011520	GGGTCAAGGCTGACCT	3597	-2430	-2500	Mb	Chr18	NT_039474	GGGTCAAGCTGATCT	-1981	-1981	-30066	5578	449	27566
MBD1	Chr8	NT_010966	GGGTCAAGGCTGACCC	-5631	-5631	-7115	Mbd1	Chr18	NT_039474	GGGTCAAGCTGATCT	-1981	-1981	-30066	5578	449	27566
MBP	Chr18	NT_025004	GGGTGCTGAGTACCC	-3322	-3399	-3409	Mbp	Chr18	NT_039478	GGGTGCTGAGTACCC	-3322	-3744	-3772	405	345	2683
MCM3AP	Chr21	NT_011515	GGGTGAGCTGACCC	6773	-7635	-7671	Mcm3ap	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	2199	3010	1929
MCM3AP	Chr21	NT_011515	GGGTGAGCTGACCC	6699	-7561	-7597	Mcm3ap	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	2125	2936	1855
MCM3AP	Chr21	NT_011515	GGGTGAGCTGACCC	6625	-7487	-7523	Mcm3ap	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	2051	2862	1781
MCM3AP	Chr21	NT_011515	GGGTGAGCTGACCC	6551	-7413	-7449	Mcm3ap	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	1977	2788	1707
MCM3AP	Chr21	NT_011515	GGGTGAGCTGACCC	6477	-7339	-7375	Mcm3ap	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	1903	2714	1633
MCP	Chr1	NT_004511	AGGTCAAGCTGACCT	-6403	-7265	-7301	Mcp	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	1829	2640	1559
MDS26	Chr10	NT_024040	AGGTCAAGCTGACCC	-1021	-7230	-11896	Mds26	Chr10	NT_039496	GGGTGCTGAGTACCT	-4574	-4625	-5742	1716	1131	3504
MFAP2	Chr1	NT_030584	AGGTCAATCTGACCC	4438	3476	1105	Mfap2	Chr4	NT_039266	GGGTCAATGATGACCT	-851	-8361	-8392	7216	131	184
MFAP2	Chr1	NT_030584	AGGTCAATCTGACCC	4438	4418	1105	Mfap2	Chr4	NT_039266	GGGTCAATGATGACCT	-851	-8361	-8392	7216	131	184
MFID	Chr3	NT_029928	AGGTCAAGTGAACCC	-1454	-1486	-1555	Mfid	Chr16	NT_039624	GGGTGCTGATGACCT	-2071	-2071	-1147	1084	1087	42
MFID	Chr3	NT_029928	AGGTCAAGTGAACCC	-1454	-1486	-1555	Mfid	Chr16	NT_039624	GGGTGCTGATGACCT	-2071	-2071	-1147	617	585	633
MFID	Chr3	NT_029928	AGGTCAAGTGAACCC	-1454	-1486	-1555	Mfid	Chr16	NT_039624	GGGTGCTGATGACCT	-2071	-2071	-1147	617	585	633
MGCC12992	Chr9	NT_021937	GGGTCAAGCTGACCT	4857	-4857	-13714	Mgcc12992	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235
MGCC2865	Chr19	NT_008470	GGGTCAAGCTGACCT	47378	1123	-8978	Mgcc2865	Chr4	NT_039268	GGGTCAAGGTTGACCC	-5208	-5413	-11479	351	556	2235

Table Supplémentaire 1-suite

RUDT2	Chr9	NT_008413	ATGTGAGTCTGACCA	2617	2617	-6917	14nd2	Chr9	NT_039259	ATGTGAGTCTGACCA	2981	2981	-9387	364	364	2470
OAZ2	Chr15	NT_010194	CGGTCAATTTGAGCTT	-1488	-1488	-1695	Oaz2	Chr9	NT_039474	AGGTCAATTTGAGCA	-1480	-1522	-1675	8	34	20
OCIA	Chr4	NT_006238	CGGTCAATTTGAGCC	25896	-10	-1547	6030432:107Ruk	Chr5	NT_039305	AGGTGCGACTGCGCT	444	392	-1045	25452	402	502
OCIA	Chr4	NT_006238	CGGTCAATTTGAGCC	25896	-10	-1547	6030432:107Ruk	Chr5	NT_039305	AGGTGCGACTGCGCT	-2049	-2101	-3538	27945	2091	1991
OFD1	Chr2X	NT_011757	AGGTGAGCTTACGCC	-6331	-6353	-6664	Ofd1	Chr2X	NT_039726	GGGTCAACCTGAGCT	-5459	-5459	-7628	872	894	964
OFD1	Chr2X	NT_011757	AGGTGAGCTTACGCC	-5631	-5653	-5364	Ofd1	Chr2X	NT_039726	GGGTCAACCTGAGCT	-5459	-5459	-7628	872	406	2264
OPV1	Chr1	NT_004836	AGGTAAAGGTGACCC	388	349	294	Opv1	Chr18	NT_039674	CAGTCAAGGTGACCC	46015	3116	-71553	84425	1687	20027
OSBPL1A	Chr8	NT_010966	AGGTCAATGATAAAGCT	130440	4803	-42326	Osbp1a	Chr18	NT_039674	CAGTCAAGGTGACCC	45799	3000	-71389	84441	1723	20963
OSBPL1A	Chr8	NT_010966	AGGTCAATGATAAAGCT	130440	4803	-42326	Osbp1a	Chr18	NT_039674	CAGTCAAGGTGACCC	45799	3000	-71389	84441	1723	20963
OSR1	Chr3	NT_037565	TGTTCAAGTTGAGCT	-2273	-2273	-2615	Osrl	Chr12	NT_039539	AGGTCAACTACACT	-3156	-7807	-7844	883	5534	5329
P2RX5	Chr17	NT_010718	AGGTCAAGCTGACCA	2754	2754	2470	P2rx5	Chr11	NT_039520	GGGTCAAGGTGACCT	4683	4683	4611	1929	1929	2141
P4HA2	Chr5	NT_034772	AGGTGCGACGACCC	67130	-1319	-9559	P4ha2	Chr11	NT_039520	GGGTCAAGGTGACCT	380	-212	-8923	66750	1107	636
p68alpha	Chr19	NT_011295	GGGTGAGCTGACCT	3218	3218	-30820	BC031407	Chr19	NT_039462	AGGTCAAGGTGACCA	82272	2145	-1101	79054	1073	29701
p68alpha	Chr19	NT_011295	GGGTGAGCTGACCT	3474	3474	-30546	BC031407	Chr19	NT_039462	AGGTCAAGGTGACCA	82272	2145	-1101	79054	1073	29701
PAD1A	Chr1	NT_030594	TGTTCAAGTGAAGCT	4744	-4744	-4770	Pads	Chr12	NT_039551	AGGTCAAGGTGACCC	-3866	-3893	-4083	1401	1374	1574
PAX9	Chr14	NT_026437	CGGTCAAGATGACTC	-5267	-5267	-5657	Pax9	Chr12	NT_039551	AGGTCAAGGTGACCC	-3292	-3319	-3509	1975	1943	2148
PAX9	Chr14	NT_026437	CGGTCAAGATGACTC	-5267	-5267	-5657	Pax9	Chr12	NT_039551	AGGTCAAGGTGACCC	-3292	-3319	-3509	1975	1943	2148
PAX9	Chr14	NT_026437	CGGTCAAGATGACTC	-5267	-5267	-5657	Pax9	Chr12	NT_039551	AGGTCAAGGTGACCC	-3292	-3319	-3509	1975	1943	2148
PAX9	Chr14	NT_026437	CGGTCAAGATGACTC	-5267	-5267	-5657	Pax9	Chr12	NT_039551	AGGTCAAGGTGACCC	-3292	-3319	-3509	1975	1943	2148
PC	Chr11	NT_033903	AGTTGCTGATGACCC	41374	-9058	-44850	Pcc	Chr19	NT_039684	ATGTGCTGTTGACCC	74771	-919	-16358	33397	61	28492
PC	Chr11	NT_033903	AGTTGCTGATGACCC	41374	-9058	-44850	Pcc	Chr19	NT_039684	ATGTGCTGTTGACCC	74771	-919	-16358	33397	61	28492
PCBP4	Chr3	NT_005999	TGTTCCCACTGACCT	-4816	-4816	-6468	Pcbp4	Chr9	NT_039477	AGGTCAAGGTGACCT	-6974	-6974	-7040	2158	581	572
PCBP4	Chr3	NT_005999	TGTTCCCACTGACCT	-4816	-4816	-6468	Pcbp4	Chr9	NT_039477	AGGTCAAGGTGACCT	-6974	-6974	-7040	2158	581	572
PCBP4	Chr3	NT_005999	TGTTCCCACTGACCT	-4816	-4816	-6468	Pcbp4	Chr9	NT_039477	AGGTCAAGGTGACCT	-6974	-6974	-7040	2158	581	572
PCDH12	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pcdh12	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PCDH8	Chr13	NT_045424	CGGTCAAGGTGACCC	3754	3754	3551	Pcdh8	Chr14	NT_039606	GTGTCAAGGTGACCT	3987	3908	3796	233	154	245
PCDH8	Chr13	NT_045424	CGGTCAAGGTGACCC	3754	3754	3551	Pcdh8	Chr14	NT_039606	GTGTCAAGGTGACCT	3987	3908	3796	233	154	245
PCMT1	Chr5	NT_023451	AGTCAAGCTGACCA	-6049	-6243	-6507	Pcmt1	Chr10	NT_039491	ATGTCAAGGTGACCA	10621	-2925	-7386	16670	831	879
PDE8A	Chr15	NT_010274	CTGTCAAGTTGACCC	-1115	-2583	-2765	Pde8a	Chr7	NT_039428	CTGTCAAGTTGACCC	-1914	-1914	-2091	799	669	674
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289	AGTCAAGTTGACCC	-4640	-4640	-5851	Pfgfrb	Chr18	NT_039674	AGGTCAAGTTGACCC	-5512	-5554	-5841	872	914	10
PFGFRB	Chr5	NT_029289</														

Table Supplémentaire 1-suite

RBN01	Chr6	NT_033951	GGGTACATGAGGACCT	-7177	-8224	-8299	Rng1	Chr17	NT_039649	AGGTACCATGAGCC	-6611	-7803	-7810	566	416	489
RBT1	Chr1	NT_004648	AGGTACATGAGGACCT	-1501	-1479	-905	Rut1	Chr3	NT_039234	TGTTACAGTTGTCT	699	647	199	802	832	706
RNF14	Chr5	NT_029289	AGGTCAAGTTGAGCC	-3321	-3321	-3321	Rra1	Chr18	NT_039674	AGGTCAAGTTGAGCC	-6663	-6671	-10419	3522	1261	532
RORA	Chr15	NT_010194	GAGTCAAGTTGAGCC	-4086	-4151	-4194	Rpa1	Chr9	NT_039474	GGGCAAGGAGGACCT	-2910	-3443	-3442	411	55	121
RPA1	Chr17	NT_010718	AGGTCAAGTTGAGCC	-3665	-3745	-3504	Rpg1	Chr11	NT_039520	AGGTAAAAGTGAAC	-4065	-4096	-4142	21	55	52
RPRR	ChrX	NT_011568	GGTCACTCTGAGCC	-5617	-5633	-5643	Rpl10a	ChrX	NT_039700	GGGTCAAGTTGAGCC	-4101	-4303	-4328	436	563	524
RPL10A	Chr6	NT_007592	AGGTGGAGATGAGCC	-8358	-8358	-8446	Rpl12	Chr2	NT_039649	AGGTCAAGTTGAGCC	-5955	-5955	-5970	338	322	322
RPL12	Chr9	NT_003470	GGGCAAGGTTGAGCC	-7335	-7294	-7316	Rpl12	Chr17	NT_039205	AGGTCAAGTTGAGCC	-7609	-7620	-7697	749	738	749
RPL22	Chr1	NT_038954	AGGTCAAGTTGAGCC	-764	-764	-794	Rpl7a	Chr4	NT_039288	AGGTCAAGTTGAGCC	-3194	-3316	-3336	10549	1022	1020
RPL7A	Chr9	NT_035014	AGGTCAAGTTGAGCC	-5317	-5348	-5400	Rps16	Chr2	NT_039205	GGGTCAAGTTGAGCC	-6828	-6830	-6882	1511	1462	1482
RPS16	Chr19	NT_011109	AGGTCAAGTTGAGCC	-7291	-7634	-7697	Rps25	Chr9	NT_039473	AGGTCAAGTTGAGCC	-8791	-8791	-8889	1500	1157	1192
RPS25	Chr12	NT_029419	GGGTCAAGTTGAGCC	-1699	-1408	-1387	Rps26	Chr10	NT_039502	GGGTCAAGTTGAGCC	-1351	-845	-845	348	563	542
RPS26	Chr1	NT_037485	GGGTCAAGTTGAGCC	-5189	-5213	-5360	Rps30a1	Chr4	NT_039266	GGGTCAAGTTGAGCC	-3447	-3447	-3591	1742	1766	1769
RPS30A1	Chr1	NT_033903	AGGTCAAGTTGAGCC	-4344	-4369	-4369	Rps30b2	Chr19	NT_039684	GGGTCAAGTTGAGCC	-4607	-4716	-4716	263	347	347
RPS30B2	Chr11	NT_033903	AGGTCAAGTTGAGCC	-4344	-4369	-4369	Rtn2	Chr19	NT_039684	GGGTCAAGTTGAGCC	-2719	-2828	-2828	1625	1541	1541
RTN2	Chr19	NT_011109	AGGTCAAGTTGAGCC	-2903	-3002	-3131	Rtn2	Chr7	NT_039400	GGGTCAAGTTGAGCC	-6975	-6975	-7140	1677	1746	1740
RTRN	Chr9	NT_011109	AGGTCAAGTTGAGCC	-4652	-4751	-4880	Rtrb	Chr7	NT_039400	GGGTCAAGTTGAGCC	-4301	-4301	-4466	1398	1209	1335
RXRb	Chr6	NT_007592	GGGTCAAGTTGAGCC	-477	-477	-456	Sap18	Chr17	NT_039649	AGGTCAAGTTGAGCC	-840	-840	-840	37	163	16
SAP18	Chr13	NT_007799	GGGTCAAGTTGAGCC	-1312	-1312	-1376	Sara	Chr14	NT_039606	AGGTCAAGTTGAGCC	-483	-446	-417	1795	1758	1793
SARA	Chr10	NT_008583	AGGTCAAGTTGAGCC	-3609	-3624	-12219	Scamp5	Chr10	NT_039495	AGGTCAAGTTGAGCC	-3950	-3996	-3561	341	372	3658
SCAMP5	Chr15	NT_010194	AGGTCAAGTTGAGCC	-1893	-1893	-267	Sgrf	Chr9	NT_039474	AGGTCAAGTTGAGCC	-283	-246	-15789	16924	711	16716
SCDF1	Chr2	NT_015926	GGGTCAAGTTGAGCC	-4092	-1012	-1258	Sgrf	Chr17	NT_039420	AGGTCAAGTTGAGCC	-2065	-2130	-2259	172	237	188
SDC1	Chr22	NT_011520	GGGTCAAGTTGAGCC	-4350	-5196	-5216	Sid1	Chr12	NT_039538	AGGTCAAGTTGAGCC	-1619	-1687	-1927	927	675	669
SDC2	Chr22	NT_011520	GGGTCAAGTTGAGCC	-2686	-1840	-1820	Sid1	Chr16	NT_039624	AGGTCAAGTTGAGCC	-1621	-1812	-1799	835	21	21
SEMA4G	Chr10	NT_030059	GGGTCAAGTTGAGCC	-3116	-109	-275	Sema4g	Chr19	NT_039692	GGGTCAAGTTGAGCC	-2915	-38	-341	201	71	66
SFPQ	Chr1	NT_004511	AGGTCAAGTTGAGCC	-7839	-7839	-7924	Sfpq	Chr11	NT_039265	GGGTCAAGTTGAGCC	-8881	-9713	-9744	1042	1874	1820
SFRS3	Chr17	NT_010461	AGGTCAAGTTGAGCC	-822	-221	-66	Sfrs3	Chr11	NT_039521	AGGTCAAGTTGAGCC	-227	-219	-66	595	2	0
SFTB	Chr2	NT_012214	AGGTCAAGTTGAGCC	-1368	-1368	-1368	Sfrs3	Chr10	NT_039350	GGGTCAAGTTGAGCC	-305	-305	-738	1673	1667	2100
SFTB	Chr19	NT_011255	GGGTCAAGTTGAGCC	-3207	-3244	-17495	Sfrs3	Chr12	NT_039496	AGGTCAAGTTGAGCC	-153	-1606	-9546	1677	1640	7949
SFTB	Chr19	NT_011255	GGGTCAAGTTGAGCC	-3207	-3244	-17495	Sfrs3	Chr16	NT_039496	AGGTCAAGTTGAGCC	-153	-1606	-9546	1677	1640	7949
SFTB	Chr19	NT_011255	GGGTCAAGTTGAGCC	-3207	-3244	-17495	Sfrs3	Chr16	NT_039496	AGGTCAAGTTGAGCC	-153	-1606	-9546	1677	1640	7949
SFTB	Chr19	NT_011255	GGGTCAAGTTGAGCC	-3207	-3244	-17495	Sfrs3	Chr16	NT_039496	AGGTCAAGTTGAGCC	-153	-1606	-9546	1677	1640	7949
SH3BP1	Chr22	NT_011520	GGGTCAAGTTGAGCC	-1214	-1214	-2415	Sh3bp1	Chr15	NT_039621	GGGTCAAGTTGAGCC	-1754	-809	-1040	2968	405	1375
SIAT6	Chr1	NT_004852	GGGTCAAGTTGAGCC	-604	-584	-28132	Siata6	Chr4	NT_039264	GGGTCAAGTTGAGCC	-426	-104	-26857	174	688	1275
SIGIRR	Chr1	NT_035113	GGGTCAAGTTGAGCC	-4186	-3244	-10767	Sigirr	Chr11	NT_039436	AGGTCAAGTTGAGCC	-9755	-9755	-9869	6509	6511	893
SIRT7	Chr17	NT_010663	GGGTCAAGTTGAGCC	-5012	-5012	-5045	Sirt7	Chr11	NT_039521	GGGTCAAGTTGAGCC	-4604	-4845	-7821	718	167	2786
SLAMF7	Chr1	NT_004648	AGGTCAAGTTGAGCC	-7077	-7243	-7377	Slamf7	Chr11	NT_039436	AGGTCAAGTTGAGCC	-9755	-9755	-9869	6509	6511	893
SLC10A2	Chr13	NT_009952	AGGTCAAGTTGAGCC	-95	-382	-979	Slc10a2	Chr11	NT_039436	AGGTCAAGTTGAGCC	-9755	-9755	-9869	6509	6511	893
SLC11A1	Chr2	NT_005403	AGGTCAAGTTGAGCC	-3175	-3937	-3937	Slc11a1	Chr1	NT_039436	AGGTCAAGTTGAGCC	-9755	-9755	-9869	6509	6511	893
SLC12A1	Chr16	NT_010498	GGGTCAAGTTGAGCC	-4545	-4526	-4520	Slc12a1	Chr8	NT_039467	GGGTCAAGTTGAGCC	-4040	-4028	-4016	505	498	504
SLC12A4	Chr16	NT_010498	GGGTCAAGTTGAGCC	-8109	-8141	-8210	Slc12a4	Chr8	NT_039467	GGGTCAAGTTGAGCC	-4040	-4028	-4016	505	498	504
SLC1A1	Chr9	NT_008413	AGGTCAAGTTGAGCC	-4654	-4699	-4890	Slc1a1	Chr9	NT_039287	GGGTCAAGTTGAGCC	-6139	-6256	-6332	1485	1557	1442
SLC25A12	Chr2	NT_005403	AGGTCAAGTTGAGCC	-8985	-8985	-8985	Slc25a12	Chr2	NT_039287	GGGTCAAGTTGAGCC	-6139	-6256	-6332	1485	1557	1442
SLC25A12	Chr2	NT_005403	AGGTCAAGTTGAGCC	-8985	-8985	-8985	Slc25a12	Chr2	NT_039287	GGGTCAAGTTGAGCC	-6139	-6256	-6332	1485	1557	1442
SLC25A18	Chr22	NT_011519	GGGTCAAGTTGAGCC	-5938	-5938	-2466	Slc25a18	Chr2	NT_039287	GGGTCAAGTTGAGCC	-6139	-6256	-6332	1485	1557	1442
SLC29A1	Chr6	NT_007592	GGGTCAAGTTGAGCC	-4315	-4242	-3416	Slc29a1	Chr17	NT_039655	AGGTCAAGTTGAGCC	-6903	-7739	-5553	2589	501	6949
SLC31A1	Chr9	NT_008470	AGGTCAAGTTGAGCC	-3773	-3773	-30814	Slc31a1	Chr4	NT_039262	GGGTCAAGTTGAGCC	-4141	-4143	-28955	7916	7916	1859
SLC35B1	Chr17	NT_010783	AGGTCAAGTTGAGCC	-3303	-3060	-2973	Ugat12	Chr11	NT_039521	GGGTCAAGTTGAGCC	-2042	-2789	-2701	261	271	272
SLC4A2	Chr7	NT_007914	GGGTCAAGTTGAGCC	-2931	-2900	-564	Slc4a2	Chr5	NT_039289	GGGTCAAGTTGAGCC	-4305	-2845	-578	1374	55	14
SLC5A5	Chr19	NT_011295	AGGTCAAGTTGAGCC	-80	-80	-427	Slc5a5	Chr5	NT_039289	GGGTCAAGTTGAGCC	-4305	-2845	-578	1374	55	14
SLU7	Chr5	NT_031313	AGGTCAAGTTGAGCC	-7785	-8092	-11825	Slu7	Chr5	NT_039289	GGGTCAAGTTGAGCC	-4305	-2845	-578	1374	55	14
SMARCB1	Chr22	NT_011520	AGGTCAAGTTGAGCC	-3246	-3306	-3453	Smarcb1	Chr10	NT_039510	GGGTCAAGTTGAGCC	-2212	-2222	-2415	1034	1034	1038
SMARCB1	Chr22	NT_011520	AGGTCAAGTTGAGCC	-3246	-3306	-3453	Smarcb1	Chr10	NT_039510	GGGTCAAGTTGAGCC	-2212	-2222	-2415	1034	1034	1038
SN	Chr20	NT_011387	AGGTCAAGTTGAGCC	-3114	-3114	-3114	Sn	Chr20	NT_011387	AGGTCAAGTTGAGCC	-3114	-3114	-3114	1034	1034	1038
SNB2	Chr16	NT_010498	GGGTCAAGTTGAGCC	-3774	-3776	-3796	Snb2	Chr16	NT_010498	GGGTCAAGTTGAGCC	-3774	-3776	-3796	1659	1659	1629
SNX4	Chr3	NT_005612	GGGTCAAGTTGAGCC	-6094	-6094	-6094	Snx4	Chr3	NT_005612	GGGTCAAGTTGAGCC	-6094	-6094	-6094	1659	1659	1629
SNR1L	Chr11	NT_033899	AGGTCAAGTTGAGCC	-3138	-3138	-3259	Snr1l	Chr16	NT_039474	AGGTCAAGTTGAGCC	-5193	-5234	-5239	811	770	807
SOX13	Chr1	NT_034410	AGGTCAAGTTGAGCC	-41828	-2027	-2027	Sox13	Chr9	NT_039474	AGGTCAAGTTGAGCC	-5193	-5234	-5239	811	770	807
SOX13	Chr1	NT_034410	AGGTCAAGTTGAGCC	-41828	-2027	-2027	Sox13	Chr9	NT_039474	AGGTCAAGTTGAGCC	-5193	-5234	-5239	811	770	807
SOX9	Chr17	NT_010641	GGGTCAAGTTGAGCC	-2691	-2691	-3063	Sox9	Chr1	NT_039521	GGGTCAAGTTGAGCC	-818	-874	-1177	1873	1817	1826
SPI	Chr12	NT_029419	GGGTCAAGTTGAGCC	-2022	-2022	-1925	Spi1	Chr15	NT_039621	AGGTCAAGTTGAGCC	-2090	-2090	-2002	68	68	77
SPAG4	Chr20	NT_028392	AGGTCAAGTTGAGCC	-3120	-3120	-3237	SPAG4	Chr2	NT_039210	AGGTCAAGTTGAGCC	-5049	-5049	-5201	1929	1929	1964
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607	-5607	-6831	Spk2	Chr7	NT_039420	AGGTCAAGTTGAGCC	-5034	-5096	-9371	573	511	2540
SPHK2	Chr19	NT_011109	GGGTCAAGTTGAGCC	-5607												

Table Supplémentaire 1-suite

TCF21	Chr6	NT_025741	ATGTCAACAGTGACCA	-8927	-8927	-9126	Tc21	Chr10	NT_039492	TGGTCACTCTGACAT	-8732	-8732	-8957	195	195	229
TCF21	Chr6	NT_025741	AGGTCACTCTGACCC	1365	1365	1106	Tc21	Chr10	NT_039492	AGGCCACCGGACCC	391	391	166	974	974	940
TCIRG1	Chr11	NT_033903	AGGTCAAGTGAACAC	-642	-642	-5215	Targ1	Chr19	NT_039684	GGGTCAAGTTGAACT	-2260	-2260	-4656	1618	2346	559
TCIRG1	Chr11	NT_033903	AGGTCAAGTGAACAC	-642	-642	-2898	Targ1	Chr19	NT_039684	AGGTCAAGGTGTCTCT	-1061	-1061	-3457	419	419	559
TCL1A	Chr14	NT_026437	GGGTCACTCGGACCT	-3299	-3299	-3384	Td1	Chr12	NT_039553	GGGTCACTGTGATCT	-5273	-5273	-5335	1974	1889	1906
TEAD2	Chr19	NT_011109	AGTTCAATCCGACCT	2970	2970	588	Tmd2	Chr7	NT_039420	AGGTGGGCGAGACCT	5167	5167	3617	2197	647	425
Tenr	Chr4	NT_016354	AGTTCAAGCTTGACCT	2133	2133	1076	Tenr	Chr7	NT_039420	TGTTCAAGCTTGACCT	1553	1553	1501	580	580	580
Tenr	Chr4	NT_016354	AGTTCAAGCTTGACCT	2133	2133	1076	Tenr	Chr3	NT_039252	TGTTCAAGCTTGACCT	1484	1484	1432	649	649	356
TERF2IP	Chr16	NT_024797	CGTTCACAAATGACCT	-616	-616	-754	Terf2ip	Chr8	NT_039467	TGGCCAATCTGACCT	526	492	415	1142	1108	1169
TESK1	Chr9	NT_008413	GGGTCAACAAATGACCT	-942	-991	-1241	Teski	Chr8	NT_039260	GGGTCAACAAATGACCT	515	-981	-981	1457	10	260
TFAP2C	Chr20	NT_011362	TGTTCAAGCTTGACCC	147	74	92	Tcfap2c	Chr2	NT_039211	TGGCCACCGTGACCC	150	129	-92	3	55	0
TFPI1	Chr21	NT_030188	GAGTCATCAGGACCT	-4524	-4562	-5660	Tfpi1	Chr8	NT_039455	GGGCCAAAATGACCA	-3095	-4018	-4072	1429	544	1588
TFPI2	Chr7	NT_007933	AGGTCCCTTGACCC	-2664	-3907	-3934	Tfpi2	Chr17	NT_039649	ATGTCAAGCTTGACCT	-4349	-4349	-4384	332	34	69
TGFBI2	Chr1	NT_077937	AGGTCAACAAATGACAC	-4370	-6816	-6992	Tgfb2	Chr5	NT_039331	AGGTCAAGTGTGACCT	-4702	-4702	-4789	2038	795	833
THBS2	Chr6	NT_078234	AGGTCAAAAGTCAACCA	-7422	-7422	-7422	Thbs2	Chr1	NT_039189	GGTTCATGGTGACCC	-6850	-6850	-8065	480	34	1073
THY1	Chr11	NT_033899	GGGTCAAGGCTGAACT	3870	2886	647	Thy1	Chr17	NT_039641	AGGTCAAGTGTGATCT	-4027	-4242	-6918	3395	3120	504
THY1	Chr11	NT_033899	GGGTCAAGGCTGAACT	3870	2886	-1571	Thy1	Chr9	NT_039473	AGTTCAAGCTTGACCC	3289	749	749	581	2137	102
THY1	Chr11	NT_033899	GGGTCAAGGCTGAACT	3870	2912	647	Thy1	Chr9	NT_039473	AGTTCAAGCTTGACCC	3289	749	749	581	2137	2320
THY1	Chr11	NT_033899	GGGTCAAGGCTGAACT	3870	2912	647	Thy1	Chr9	NT_039473	AGTTCAAGCTTGACCC	3289	749	749	581	2163	102
TIMELESS	Chr12	NT_029419	ATGTCAAGGCTGACCC	-1039	-1039	-16252	Timeless	Chr9	NT_039473	AGTTCAAGCTTGACCC	3289	749	749	581	2163	2320
TIRAP	Chr11	NT_033899	TGTTCAAGATGACCT	-7826	-7826	-7826	Tirap	Chr10	NT_039502	AGGTCACTCTGACCT	-2617	-8626	-10462	1578	7587	5790
TIRAP	Chr11	NT_033899	TGTTCAAGATGACCT	-7826	-7826	-15555	Tirap	Chr9	NT_039473	AGTTCAAGCTTGACCC	3289	749	749	581	2163	2320
TIRAP	Chr11	NT_033899	TGTTCAAGATGACCT	-7826	-7826	-15555	Tirap	Chr9	NT_039473	AGTTCAAGCTTGACCC	3289	749	749	581	2163	2320
TIP3	Chr19	NT_011255	GGGTCAATAGTGACCC	20860	829	829	Tip3	Chr10	NT_039496	GGGTCAAGCTGACCC	3244	2897	-4606	17616	2088	5435
TIP3	Chr19	NT_011255	GGGTCAATAGTGACCC	20860	829	829	Tip3	Chr10	NT_039496	GGGTCAAGCTGACCC	3244	2897	-4606	17616	2088	5435
TLX1	Chr10	NT_030059	CTGTCAAGGTGACCC	-5805	-5805	-6842	Tlx1	Chr19	NT_039692	GGGCCAGGCGGACCT	-7158	-7359	-7359	1353	1554	517
TMPO	Chr12	NT_019546	AGGTCTTTATGACCT	-7368	-7439	-7643	Tmpos	Chr10	NT_039500	CTGTCAAGCTGACCT	-6064	-6081	-6303	1304	1358	1340
TMPS56	Chr22	NT_011520	GGGACGCTCTGACCC	1087	1087	16770964	Tmpos56-pending	Chr15	NT_039621	AGGTCACTCTGACCC	303	155	-2982	784	932	16773946
TNFRSF1B	Chr17	NT_010718	AGGTCAAGCTGACCT	4550	4550	4537	Tnfrsf1b	Chr15	NT_039621	AGGTCACTCTGACCC	303	155	-2982	784	932	16773946
TNFRSF9	Chr1	NT_020554	AGTTCAAGATGACCT	-2848	-2848	-3681	Tnfrsf9	Chr15	NT_039621	AGGTCACTCTGACCC	303	155	-2982	784	932	16773946
TNP1	Chr2	NT_005403	AGGTCAATATGCCCA	-6888	-6909	-6913	Tnp1	Chr4	NT_039268	TGTTCAATGTGTGACCC	10938	194	4346	13786	3042	665
TNP1	Chr2	NT_005403	AGGTCAATATGCCCA	-6888	-6909	-6913	Tnp1	Chr1	NT_039170	CTGTCAAGCTTGACCC	-7294	-7315	-7324	406	406	411
TP53	Chr17	NT_010718	AGGTCAAGCTGACCT	-1529	-1529	-12534	Tp53	Chr1	NT_039170	CTGTCAAGCTTGACCC	-7294	-7315	-7324	406	406	411
TP53BP1	Chr15	NT_010194	AGGCCAGGCTGACCA	-2590	-2590	-20668	Tp53bp1	Chr11	NT_039520	TGGGCACTTGACCC	-2046	-2046	-8956	517	517	3578
TPB3	Chr6	NT_007299	TGTTCAAGCTGACCT	1876	820	736	Tpb3	Chr2	NT_039209	TGGACATTTGACCT	-4095	-4095	-4186	1505	1505	16482
TPH1	Chr12	NT_009759	AGGTCTCTGACCC	2578	2169	2142	Tph1	Chr17	NT_039475	TGTTCAAGCTGACCT	2324	2324	754	448	1504	18
TPH1	Chr12	NT_009759	AGGTCTCTGACCC	2578	2169	2142	Tph1	Chr6	NT_039369	GAGTCACTGACCC	2021	2021	1731	557	148	41
TPM3	Chr1	NT_004668	GGGTCAAGCTGACCC	1419	1387	1384	Tpm3	Chr6	NT_039369	GAGTCACTGACCC	2021	2021	1731	557	148	41
TPM3	Chr1	NT_004668	GGGTCAAGCTGACCC	1419	1387	1384	Tpm3	Chr6	NT_039369	GAGTCACTGACCC	2021	2021	1731	557	148	41
TPS1	Chr16	NT_037887	AGGTCAAGCTGACCC	-3630	-3630	-3892	Tps1	Chr3	NT_039356	GGGTCAAGTGAACCT	2105	2105	1815	473	64	327
TPS1	Chr16	NT_037887	AGGTCAAGCTGACCC	-3630	-3630	-3892	Tps1	Chr3	NT_039356	GGGTCAAGTGAACCT	2105	2105	1815	473	64	327
TRAF3	Chr14	NT_026437	AGGTCAAGCTGACCA	-7549	-7549	-100772	Traf3	Chr3	NT_039356	GGGTCAAGTGAACCT	2105	2105	1815	473	64	327
TRAF3	Chr14	NT_026437	AGGTCAAGCTGACCA	-7549	-7549	-100772	Traf3	Chr3	NT_039356	GGGTCAAGTGAACCT	2105	2105	1815	473	64	327
TRAP25	Chr8	NT_008046	TGTTCACTCTGATCC	-8316	-8394	-8420	Trap25	Chr15	NT_039621	AGGTCACTCTGACCC	-9127	-9127	-9192	811	733	712
TREX1	Chr3	NT_005825	TGTTCAAGCTGACCC	-8074	-8157	-8180	Trex1	Chr15	NT_039621	AGGTCACTCTGACCC	-9127	-9127	-9192	811	733	712
TRIM11	Chr1	NT_004559	GGGTCAAGCTGACCC	7738	-7763	-8018	Trim11	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
TRIM11	Chr1	NT_004559	GGGTCAAGCTGACCC	7738	-7763	-8018	Trim11	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
TRIM26	Chr6	NT_034574	TGTTCAAGCTGACCA	7889	-696	-6359	Trim26	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
TRIM26	Chr6	NT_034574	TGTTCAAGCTGACCA	7889	-696	-6359	Trim26	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
TRIM41	Chr5	NT_077451	AGGTCAAGCTGACCC	2215	-3157	-3180	Trim41	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TRIM41	Chr5	NT_077451	AGGTCAAGCTGACCC	2215	-3157	-3180	Trim41	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TRPM1	Chr15	NT_010194	AGGTCAAGCTGACCC	3787	2845	2822	Tpm1	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TRPM2	Chr21	NT_011515	GGGTCTTTTGACCT	3790	3790	3345	Tpm2	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TRPM5	Chr11	NT_023310	GGGTCTTTTGACCT	958	958	949	Tpm5	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TSB	Chr1	NT_019273	AGGTCAAGCTGACCT	-577	-4099	-4099	Tsb	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TSB	Chr1	NT_019273	AGGTCAAGCTGACCT	-577	-4099	-4099	Tsb	Chr17	NT_039661	AGGTCACTCTGACCC	-2661	-2661	-59163	10550	1965	52804
TSDRD2	Chr22	NT_011519	AGGTCAAGCTGACCC	8508	4695	1115	Tsdrd2	Chr15	NT_039621	TGTTCACTCTGACCC	-7323	-7888	-7902	2419	2229	1545
TSDRD2	Chr22	NT_011519	AGGTCAAGCTGACCC	8508	4695	1115	Tsdrd2	Chr15	NT_039621	TGTTCACTCTGACCC	-7323	-7888	-7902	2419	2229	1545
U2AF1	Chr12	NT_030188	AGGTCAAGCTGACCC	-4824	-4887	-4917	U2af1	Chr16	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UBE2R2	Chr9	NT_008413	AGGTCAAGCTGACCC	994	755	564	Ube2r2	Chr16	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UBL4	ChrX	NT_025965	GGGTCAAGCTGACCC	4539	4462	4462	Ubl4	Chr16	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UBL5	Chr19	NT_011295	AGGTCAAGCTGACCC	-4458	-4497	-4897	Ubl5	Chr16	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UBTF	Chr16	NT_010393	GGGTCAAGCTGACCC	-2487	-2487	-2487	Ubtf	Chr16	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UBTF	Chr16	NT_010393	GGGTCAAGCTGACCC	-2487	-2487	-2487	Ubtf	Chr16	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UCK119	Chr17	NT_010799	GGGTCAAGCTGACCC	4235	4235	4167	Uck119	Chr17	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
UCK119	Chr17	NT_010799	GGGTCAAGCTGACCC	4235	4235	4167	Uck119	Chr17	NT_039624	GGGTCAATGTGACCA	3533	3533	3505	4975	1162	4620
USF1	Chr1	NT_004668	AGGTCAAGCTGACCC	7888	-696	-6359	Usf1	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
USH1C	Chr11	NT_009237	GGGTCAAGCTGACCC	-9078	-9078	-9173	Ush1c	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
USP14	Chr18	NT_010859	GGGTCAAGCTGACCC	2477	2444	2298	Usp14	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
USP2	Chr11	NT_033899	AGGTCAAGCTGACCC	17001	-528	-715	Usp2	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
USP21	Chr1	NT_004668	TGTTCAATGTGACCT	2595	2451	1423	Usp21	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
USP21	Chr1	NT_004668	TGTTCAATGTGACCT	2595	2451	1423	Usp21	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778	1759	1890
VAT1	Chr17	NT_010755	GGGTCAAGCTGACCC	-3629	-3686	-3742	Vat1	Chr11	NT_039520	AGGTCACTCTGACCC	-5960	-6004	-6128	1778		

Chapitre 6 : Mechanismes of Gene Regulation by Estrogen Receptor Alpha in Breast Cancer Cells

Véronique Bourdeau^{1§}, Julie Deschênes^{1§}, David Laperrière¹, Malika Aid¹, John H. White² and Sylvie Mader^{12*}

[§] These authors contributed equally to this manuscript

Article soumis : *Nucleic Acid Research*

À titre de co-premier auteure, j'ai réalisé le microarrays et les différentes expériences biologiques en collaboration étroite avec Dr Bourdeau. J'ai également participé à la rédaction du manuscrit, plus particulièrement dans le montage des figures et dans ses révisions.

1 Institute for Research in Immunology and Cancer and Biochemistry Department, Université de Montréal, C.P. 6128 Succursale Centre Ville, Montréal QC H3C 3J7 Canada.

2 Departments of Physiology and Medicine, McGill University, Montréal, QC H3A 1A4 Canada.

* To whom correspondence should be addressed at: Institute for Research in Immunology and Cancer, Faculté de Médecine, Université de Montréal, C.P.6128 Succursale Centre Ville, Montréal Qc H3C 3J7 Canada. Tel: 514-3437166, Fax: 514-3436843, E-mail:

[REDACTED]

ABSTRACT

Estrogen receptors (ERs), which mediate the proliferative action of estrogens in breast cancer cells, are ligand-dependent transcription factors that regulate expression of their primary target genes through several mechanisms. In addition to direct binding to cognate DNA sequences, ERs can be recruited to DNA through other transcription factors (tethering), or affect gene transcription through modulation of signaling cascades by non-genomic mechanisms of action. To better characterize the mechanisms of gene regulation by estrogens, we have identified more than 700 putative primary and more than 1500 putative secondary target genes of estradiol in MCF7 cells through microarray analysis performed in the presence or absence of the translation inhibitor cycloheximide. Although siRNA-mediated inhibition of ER α expression antagonized the effects of estradiol on up- and down-regulated primary target genes, estrogen response elements (EREs) were enriched only in the vicinity of up-regulated genes. Binding sites for several other transcription factors, including proteins known to tether ER α , were enriched in up- and/or down-regulated primary targets. Secondary estrogen targets were particularly enriched in sites for E2F family members, several of which were transcriptionally regulated by

estradiol, consistent with a major role of these factors in mediating the effects of estrogens on gene expression and cellular growth.

INTRODUCTION

The pleiotropic effects of estrogens in its numerous target tissues, including the reproductive, skeletal, cardiovascular and central nervous systems (1-5) are mediated in large part via ERs (6), which are members of the superfamily of nuclear receptors and function as hormone-dependent transcription factors (7-9). ERs bind DNA directly through their central, conserved DNA binding domains composed of two zinc fingers of the C4 type (10,11). Cognate DNA binding motifs, also called estrogen response elements (EREs), have been characterized in estrogen responsive promoters (12-15). Consensus EREs derived by compiling natural response elements are 15 bp palindromes of PuGGTCA motifs and correspond to the highest affinity binding sites for ERs *in vitro* (16,17). However, natural response elements often deviate from the consensus at one or several positions (14,15). Both estrogen receptors share similar *in vitro* DNA binding patterns (18), but their transcriptional activation properties differ (6,19,20), possibly due to differential recruitment of transcriptional coactivator complexes responsible for histone acetylation, chromatin remodeling, and enhanced recruitment of the basal transcription machinery (21-27). ER α is thought to mediate the proliferative effects of estrogens in breast cancer cells. Indeed, its expression is preserved or increased in two thirds of breast tumors, correlating with sensitivity of tumors to antiestrogenic treatment (3-5). On the other hand, the expression of ER β , which was reported to play a role in terminal differentiation of breast epithelial cells (28), appears to be reduced during tumorigenesis (29,30).

In addition to mediating gene regulation through direct binding to DNA, ERs can regulate gene expression through protein-protein interaction with other transcription factors. Several transcription factors were shown to mediate positive or negative transcriptional regulation by ERs in the absence of EREs, including AP1, Sp1, NF κ B

(15,31-33). In addition, interference between estrogen signaling and other intracellular signaling pathways including the MAPK and PI3K pathways have been widely reported and may result from interactions between ERs and components of these signaling cascades (15,34-38). Finally, it has recently been suggested that estrogens may act also through a membrane receptor member of the GPCR family, GPR30, although the importance of these receptors in breast tumorigenesis remains to be established (39-43).

Better understanding of the mechanisms of action of estrogens in breast tumorigenesis necessitates large-scale identification of estrogen target genes and comprehensive analysis of the mechanisms of target gene regulation to assess the contribution of different regulatory mechanisms and specific targets to the proliferative effects of estrogens. Genome-wide microarray analysis of estradiol (E2) target genes has been performed in ER α -positive breast cancer cell lines such as MCF7, T47D and ZR75 cells, leading to the identification of a large number of target genes (44-50). It is however not always clear to which degree target identification is affected by cell culture conditions, choice of microarray platform and statistical analysis tools. In addition, few studies have used conditions that distinguish between primary and secondary target genes. This may explain why enrichment in EREs in estrogen target genes identified through microarray analysis was not reported in most studies. On the other hand, the promoter regions of 89 E2 target genes regulated in the presence of CHX in T47D were found enriched in EREs (46). However, the number of identified primary E2 target genes in that study remains low and the question of how far from the transcriptional start sites of primary target genes enrichment in EREs can be observed remains open given the relatively narrow window used in this study. High-affinity EREs located distally from the start sites of estrogen target genes are functional ER-binding sites in chromatin immunoprecipitation (ChIP) experiments (50,51). Although the role of ERs bound to distal sites in transcriptional regulation remains to be systematically analyzed, chromatin conformation capture assays revealed that ER α -bound chromatin regions can act at large distances from regulated genes (48,52). In addition to EREs, binding sites for several other transcription factors were

enriched in ER α -bound ChIP fragments, suggesting roles for these factors in ER α tethering (48).

In order to better study the mechanisms of target gene regulation by estrogens, we aimed to discriminate between primary and secondary E2 target genes in MCF7 cells through microarray analysis of gene expression patterns in response to E2 either in the presence or the absence of the protein synthesis inhibitor cycloheximide. We verified the validity of this approach by demonstrating enrichment in EREs in primary, but not secondary up-regulated genes, and identified other transcription factors involved in the regulation of primary and/or secondary estrogen targets. Our results are discussed in the context of previous studies based on large-scale microarray analysis or genome-wide ChIP-on-chip mapping of ER α -associated chromatin regions.

MATERIALS AND METHODS

Cell culture and treatments

MCF-7 breast carcinoma cells were maintained in α -minimal Eagle's medium (α -MEM) (Wisent, St-Bruno, QC, Canada) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, Oakville, On, Canada). Three days before experiments, cells were switched to phenol red-free Dulbecco's modified Eagle's medium (DMEM) (Wisent) containing 10% charcoal-treated FBS, 1% sodium pyruvate (Wisent), 1% penicillin/streptomycin (Wisent) and 1% L-glutamine (Wisent). Cells were seeded in 10 cm plates at a density such that near-confluency is obtained at the end of the treatment. The day before hormonal stimulation, the medium was changed to phenol red-free DMEM supplemented with 0,5% charcoal-treated FBS. Pre-treatments with cycloheximide (10 μ g/ml, Sigma) were initiated 1 h before hormonal treatment with 17 β -estradiol (E2, 25 nM, Sigma), ICI 182,780 (ICI, 100 nM) or vehicle (0.1% ethanol) for variable periods of time as indicated in figure legends. Selected siGenome siRNAs (Dharmacon, Chicago, IL, USA) were transfected using siLentFect lipid reagent according to the instruction of the

manufacturer (Bio-Rad Laboratories, Hercules, CA, USA). Medium was changed 24 h after transfection with fresh phenol red-free DMEM supplemented with 0.5% charcoal-treated FBS and hormonal treatments were initiated 48 h after transfection. Cells were harvested 24 h after (i.e. 72 h after siRNA transfection).

Western analysis

Whole cell extracts were prepared as described previously (53) using antibodies directed against ER α (mouse monoclonal B10 antibody, a kind gift from Prof. P. Chambon) or against β -actin mouse monoclonal antibody (AC-15, Sigma Diagnostics).

Growth assays

MCF7 breast cancer cells were seeded to low density and treated every two to three days with vehicle (0), 17 β -estradiol (E2, 25 nM) or ICI 182-720 (ICI, 100 nM) in 5% charcoal-stripped FBS. Protein concentrations were measured after nine days as previously described (54). When siRNA were used, cells were re-plated 24 h after transfection to low density and treated twice over a 5-day period with vehicle (0), 17 β -estradiol (E2, 25 nM) or ICI 182-720 (ICI, 100 nM) prior to quantification of total protein concentrations. Experiments shown are representative of two to three independent experiments performed with duplicate samples.

Cell cycle analysis

MCF7 cells were transfected as previously described. After 24 h, medium was changed for phenol-red free DMEM supplemented with 0.5% charcoal-treated FBS and 24 h later, stimulated with either vehicle (0, 0.1% ethanol) or estradiol (E2, 25 nM) for a last 24 h. Cell cycle values were obtained after ethanol fixation and propidium iodide staining followed by data acquisition on a FACS "BD LSR II" and analysis using the ModFit LT cell cycle software. Experiments were performed at least twice with duplicate samples.

RNA purification and DNA microarrays

After completion of hormonal treatments, medium was completely removed and cells were collected in 1 ml of TRI-Reagent (Sigma). Total RNA was extracted as recommended by the manufacturer and further purified with RNeasy MinElute Cleanup Kit (QIAGEN, Mississauga, On, Canada). cDNA synthesis from total RNA, labeling and hybridization to Affymetrix HG-U133 2.0 Plus gene chips were performed at the Genome Quebec and McGill University Innovation Center using standard protocols (<http://www.genomequebec.mcgill.ca/>). Microarrays were analyzed with the affyGUI Bioconductor package (55) after normalization of probe-level data with RMA (56). Genes deemed significantly regulated were those with ≥ 1.4 fold change between vehicle and E2 treatments, average \log_2 -expression levels greater than 5 across all samples (A-value) and a p-value for moderated t-statistics smaller than 0.01. Regulated genes were ranked according to $|\text{fold}| * A * (1 - p\text{-value})$.

Screening for transcription factor binding sites

The reference human genome sequences (hg17, May 2004) were downloaded from the UCSC Genome Browser database (57). Genomic sequences in 10 kb windows centered around the Transcription Start Sites (TSS) of all annotated gene in the RefSeq Genes track (58) were extracted from the genome. These sequences were screened with 299 matrices for transcription factor binding sites using a base score cut-off of 65% and 5% increments. The program used for screening implements the formulas described by Wasserman and Sandelin (59); other programs developed for genome-wide screen of nuclear receptor response elements (51,60,61) were used to calculate the distance from the TSS and identify binding sites within transposable elements or regions bound in ChIP-on-chip experiments. The majority of matrices used was from the public version of TRANSFAC (62) or represented variation of TRANSFAC matrices (i.e. E2F matrix M00516 where only the portion matching TTTSGCGC was used). Selected matrices were compiled from the literature including matrices for estrogen response elements (51,63,64) and RXR/RAR (65). Searches were also performed with the same sequences from which transposable elements

were removed, or with only sequences present in conserved regions, as described in the “PhastCons Conserved Elements” track from the UCSC (66), within the same 10 kb windows. For each transcription factor, four cutoffs were chosen with frequencies in all gene promoters closest to those of EREs (70%: 0.32 ERE/gene, 75%: 0.16 ERE/gene, 80%: 0.06 ERE/gene, 85%: 0.02 ERE/gene). Cutoffs were required to be between 2 sites/gene and 0.01 site/gene to avoid overly abundant sites or rare sites. Z-scores and p-values were determined for each of the selected cutoffs to assess the statistical significance of the observed enrichment in promoters of regulated genes on the microarray versus those of all annotated genes. The Z-score and p-values were calculated with the oPOSSUM perl application programming interface (API) using the cutoff recommended by the authors (67). The ER α ChIP-on-chip dataset (50) was downloaded from the Brown Lab website (<http://research.dfci.harvard.edu/brownlab/datasets/index.php>) and the E2F1 ChIP-on-chip dataset (68) was downloaded from the NCBI (GSE5175).

Gene expression quantification

Total RNAs were extracted as described above and aliquots of 2 μ g were reverse transcribed using the RevertAid H first minus strand cDNA synthesis kit (MBI Fermentas, Burlington, On, Canada) as recommended by the manufacturer. Reverse transcription products were diluted 10 times in pure water prior to real-time quantitative PCR. Gene expression levels were determined using primer and probe sets from the Universal Probe Library (<https://www.roche-applied-science.com>). PCR reactions were performed in 384 well plates using 2 μ l of cDNA sample, 5 μ l of TaqMan PCR Master Mix (Applied Biosystems, CA), 2 μ M of each primer (sequences available upon request) and 1 μ M of the Universal TaqMan probe in a total volume of 10 μ l.

The ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) was used to detect amplification levels and was programmed for an initial step at 95°C for 10 min, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. All reactions were run in duplicates or triplicates and average values were calculated. Quantification was performed with at least two independent experiments. The house keeping RPLP0

(ribosomal protein, large, P0) gene was used as endogenous control. Relative expression levels of target genes and standard deviation values were determined using the $\Delta\Delta CT$ method.

Chromatin immunoprecipitation

For chromatin immunoprecipitation (ChIP) assays, chromatin was cross-linked by treating cells with 1.5% formaldehyde for 10 min at room temperature and fragmented by sonication as previously reported (51,52), yielding fragments of average size approximately 400 bp. Immunoprecipitation were conducted as previously described (51,52). Antibodies against a C-terminal epitope of hER α (SC-543) and against β -actin (SC-8432) were purchased from Santa Cruz Biotechnology, Inc (Santa Cruz, CA, USA). Detailed conditions for immunoprecipitation, PCR amplification and primer sequences (synthesized by Sigma Genosis) used in ChIP assays are available upon request. Chromatin immunoprecipitation experiments were performed twice with similar results. A representative set of results is shown.

RESULTS

Identification of primary target genes of estrogen receptor alpha in MCF7 cells.

MCF7 cells are one of the most widely used cell models to study estrogen signaling as they are strongly growth-stimulated by 17 β -estradiol (E2) *in vitro* (see Fig. 1A) and in xenograft experiments (69), and express high levels of estrogen receptor alpha (ER α). Transfection of siRNAs directed against ER α , but not GAPDH or Luciferase, strongly suppressed MCF7 cell growth (Fig. 1B), indicating that ER α expression levels are rate-limiting for growth of these cells. In the absence of E2, MCF7 cells cultured in 0.5% serum were mostly arrested in the G0-G1 phase (>90%), while addition of E2 for 24 h augmented significantly the proportion of cells in S phase (Fig. 1C). On the other hand, transfection of siRNAs against ER α reduced this proportion both under basal and induced conditions. Thus, it is likely that genes regulated by estrogens under these conditions include a large

proportion of cell cycle regulated genes, some of which may not be primarily regulated by estrogens. In addition, E2 treatment leads to regulation of transcription factors, cofactors and signaling proteins, which amplifies the genomic effects of estrogens, complicating the analysis of binding sites in the regulatory sequences of target genes.

To identify primary estrogen target genes in MCF7 cells, we performed treatments with E2 for 24 h with or without pre-treatment of cells with the protein synthesis inhibitor cycloheximide (CHX). Gene expression patterns were analyzed using Affymetrix GeneChips HG-U133 2.0 Plus chips representing over 47,000 human transcripts. Differential expression was assessed as the ratio between the average of the four duplicates in the presence of estrogen versus its absence, either in the presence of CHX or in its absence. A cut-off of 1.4 was applied based on observed regulation of known ER targets such as TFF1 under these conditions. A total of 2144 significantly regulated genes were identified and ranked according to their combined p-value, fold-regulation and intensity of expression (please see the lists of the top 100 genes regulated in the absence or presence of CHX in Table 1). There was a significant overlap in regulated genes with other microarray studies (440/1483 with Carroll et al., 2006 (50) and 361/961 with Rae et al., 2005 (70); see Fig. 2A) and the large majority (>95%) of genes in the overlapping sets were found to be regulated similarly in pair-wise comparisons. This partial overlap may reflect differences in experimental conditions (e.g. time of treatment, culture medium), in the type of microarrays (Affymetrix U133A versus U133 Plus 2.0), and/or in filtering parameters. Note however that the percentage of genes in common between the previous two studies is lower than with ours, indicating that partial overlap is not due to a bias specific to our study. Only 128 target genes were identified as estrogen target genes in all three studies.

Regulated genes included nearly equivalent numbers of up- and down-regulated genes in the absence of CHX (858 up-regulated and 838 down-regulated genes, Fig. 2B), but more than twice as many up- versus down-regulated genes in the presence of CHX (544 up-regulated and 236 down-regulated genes, Fig. 2B). Only 25 genes had opposite

regulation in the absence or presence of CHX. The proportion of genes that were not identified as E2 targets in the two previous studies large-scale studies in MCF7 cells (50,70) is similar for genes regulated in the absence (68%) and in the presence (71%) of CHX, indicating that treatment with the protein synthesis inhibitor is unlikely to account for the partial overlap with these studies. Target gene regulation was confirmed using Q-PCR for 41 putative primary and 22 putative secondary target genes (supplementary Fig. 1).

Surprisingly, the overlap between genes regulated in the presence and absence of CHX was only partial (35% for up-, 50% for down-regulated genes). This may be due to several factors, including regulation below the cut-off and/or lack of detection of transcripts in the -CHX category, or suppression of negative feedback mechanisms on ER-mediated gene expression by CHX treatment. In addition, we found that several well-known estrogen target genes were represented in the set regulated only in the presence of CHX, including CCND1, IGFBP4, C3 and MYC. Regulation by E2 of several genes in this category could be detected at earlier times (peak around 8 h after E2 stimulation, Fig. 3B), and was also observed in the presence of CHX, confirming their status as primary target genes. On the other hand, CHX treatment did not increase regulation of the immediate early FOS gene, which has a peak of expression at 1 h post E2 treatment (Fig. 4A).

Representation of genes involved in cell cycle control, including cyclins and other cell cycle regulators, is strikingly different in genes regulated in the absence or presence of CHX (data not shown). This correlates with the fact that CHX treatment led to loss of progression into S phase (supplementary Fig. 2). The group of genes whose induction by E2 was repressed by CHX includes genes associated with the proliferation signature of breast cancer cells, such as members of the E2F family of transcription factors (Fig. 4B), of the cell division cycle (CDC) and minichromosome maintenance (MCM) gene families (supplementary Fig. 1).

ER α participates in both up- and down-regulation of gene expression, but EREs are enriched only in the flanking sequences of up-regulated primary target genes.

The observation that depletion of ER α reduced MCF7 cell growth both and decreased the fraction of cells in S phase (Fig. 1B-C and supplementary Fig. 2) suggests that primary E2 target genes mediating the proliferative action of this hormone are likely under regulation by ER α . To investigate whether ER α plays a similar role in the up- and down-regulation of primary target genes, we selected a sample of 5 primary up- and 5 down-regulated target genes and assessed the effect of siRNA-mediated ER α depletion (Fig. 5). Si-RNAs directed against ER α markedly reduced levels of expression of the receptor, and levels of cyclin D1 (CCND1) mRNAs (Fig. 5A-B). Modulation of both up- (Fig. 5C) or down-regulated (Fig. 5D) primary estrogen target genes was markedly attenuated by depletion of ER α , but not of the GAPDH or LUC controls, indicating that expression levels of the receptor are a limiting factor for overall target gene regulation in MCF7 cells. These results also suggest that overexpression of ER α during tumorigenesis likely plays an important role in potentiating estrogen signaling, and that lower levels of expression of ER α in some cell lines may abrogate the regulation of the weakest target genes.

The classical mechanism of action of ER α signaling involves direct binding to cognate DNA sequences, the estrogen response elements, whose consensus sequence is a palindrome of PuGGTCA motifs (13,16,17). Surprisingly, there was no difference in the percentage of EREs identified between groups of up-, down-, or non-regulated genes in ZR75 cells (47). On the other hand, enrichment in EREs was reported in the promoter sequences of target genes identified in the presence of cycloheximide in T47D cells (46); however the number of primary target genes identified was small (89) and enrichment was assessed only in a 3.5 kb window of regulatory sequences. Here, we compared ERE frequencies in genes regulated in the presence and/or absence of the protein synthesis inhibitor cycloheximide (CHX) in MCF7 cells. EREs were identified at variable distances

from the transcriptional start sites of genes using a customized position weight matrix, derived by compiling binding motifs from the literature and by symmetrical complementation reflecting the palindromic nature of this site (Fig. 6A; note that this matrix yielded higher enrichment in EREs than the TRANSFAC matrix in control datasets corresponding to lists of genes with known EREs or ER α -bound ChIP fragments, data not shown). Windows of 20 kb centered on the transcriptional start sites (TSS) of target genes were used with different cut-offs corresponding to frequencies of sites in the genome comprised between 0.1 and 2 sites per gene. This matrix detected a significant enrichment in EREs within 20 kb of flanking sequences in a test list of genes found to contain EREs in a genome-wide search for high affinity EREs, compared to the whole genome (51). The ratio between the number of sites detected in regulated genes versus all genes in the genome remained close to one at all cut-offs when both up- and down-regulated genes were considered together, but increased with higher cut-offs in up-regulated genes while it decreased in down-regulated genes (Fig. 6B). However statistical significance ($p < 0.01$ and Z score > 10) was reached only at one point in total up-regulated genes in this window. On the other hand, ERE enrichments in E2 target genes up-regulated in the presence of CHX (+CHX group), but not in its absence (-CHX group), were statistically significant. Further subdivision of the +CHX group into only+CHX and -and+CHX indicated statistically significant enrichment in both the only+CHX group and the -and+CHX groups. No enrichment was seen in the only-CHX group, consistent with the interpretation that most of the genes in this group are secondary target genes. The observation that enrichment was higher with higher cut-offs indicates that sequences closer to the consensus are more likely to mediate transcriptional regulation, suggesting a parallel between *in vitro* affinity and *in vivo* binding.

Distribution of EREs in sliding 2.5 kb windows with 500 bp increments between -25 and +25 kb of the TSS indicates that enrichment of EREs (identified at a 75% cut-off) in genes up-regulated in the presence of CHX can be observed within -20 to +20 kb of the TSS (Fig. 7A). On the other hand, no enrichment compared to the genomic distribution was

observed when considering genes up-regulated in the absence of CHX (-CHX or only-CHX groups, Fig. 6B). Finally, enrichment in EREs was higher when considering only ER α -associated chromatin regions (50) found within a 20 kb window around the TSS of different groups of regulated genes, but was still not observed in the only-CHX group of regulated genes (Fig. 6C). Together, these results validate the use of CHX to identify primary E2 target genes.

Surprisingly, no enrichment but rather a statistically significant depletion in EREs was observed in a 20 kb region around the TSS of genes repressed by estrogens (Fig. 6B, Fig. 7B). We examined whether enrichment in EREs was observed in ER α -associated chromatin regions (71) found near genes down-regulated in the absence or presence of CHX (Fig. 6C). No enrichment in ERE sequences was observed in the genes down-regulated either in the presence or the absence of CHX. This suggests that binding sites for other transcription factors mediate the down-regulation of genes by E2 in the presence of CHX (see also next section).

To test whether predicted EREs can bind ER α *in vivo*, selected regions containing EREs identified using a 65% cut-off and present at +/- 25 kb of the TSS of target genes up-regulated in the presence of CHX were amplified by PCR after ChIP. Binding of ER α was observed on all predicted elements with the exception of the ERE with 2 mismatches present at +2.3 kb of the EGR3 TSS (Fig. 8). Only two of the nine validated elements (the WISP2 -452 ERE and CA12 -6 kb ERE) were detected previously in genome-scale ChIP experiments (71). The intensity of binding to individual elements did not necessarily reflect their degree of similarity to the consensus sequence, as one of the strongest sites bound (at -452 bp of the WISP2 TSS, also observed in ref. (71)), diverges from the consensus ERE by two mismatches. This may be due to the sequence context of the EREs, which is likely to affect *in vivo* affinity through cooperativity with other transcription factors bound close to the EREs. In addition, the occurrence of several EREs in the flanking regions of primary target genes suggests potential cooperativity between EREs for DNA binding and

transcriptional activation through formation of chromatin loops as described for the GREB1 gene (52).

Binding sites for other transcription factors are enriched in up- and down-regulated primary estrogen target genes.

The lack of enrichment in EREs in down-regulated target genes suggests transcriptional regulation via other mechanisms than ER α binding to EREs. Several transcription factors have been reported to mediate indirect recruitment of ER α to DNA via protein-protein interactions (see (15,31-33) for reviews). In addition, other transcription factors may also mediate estrogenic regulation via non-genomic mechanisms (15,34-38). Binding sites for transcription factors participating in either mechanism may be enriched in the flanking sequences of primary regulated genes, while only enrichment in tethering transcription factors is expected in ER α -associated regions identified through ChIP.

To identify transcription factor binding sites enriched in either up- or down-regulated primary targets, we screened TRANSFAC matrices for >1.5 fold enrichment in frequencies within 5 or 10 kb windows around the TSS of different classes of regulated genes versus all annotated genes. Due to the variable representation of different transcription factor binding sites in low complexity regions, this analysis was performed considering either whole genomic sequences, sequences without repeated elements or only conserved sequences. Enrichments with Z scores higher than 10 and p-values lower than 0.01 for at least two different cut-offs of the same matrix were retained as significant. This analysis was performed using either all significantly regulated genes (Supplementary Table 1) or the top 100 regulated genes (Supplementary Table 2) in each category.

Binding sites enriched in primary up-regulated target genes (regulated in the presence of CHX) included, in addition to EREs, sites for AHR/ARNT, AP1, CCAAT box factors, C/EBP homologous protein (CHOP), CHX10 homeobox factor, EGR factors, Maf,

GCbox factors and ERE-like sequences with different arrangements of the motifs (Supplementary Table 1 & 2). On the other hand, binding sites enriched in primary down-regulated genes include sites for the forkhead transcription factor FREAC3 (FOXC1), CHX10 and PAX homeobox factors, EVI1, Ikaros and SRY transcription factors (Supplementary Table 1 & 2).

E2F transcription factors are a major relay for transcriptional regulation by estrogens of secondary target genes in MCF7 cells.

No enrichment of EREs was observed in genes regulated in the absence of CHX only, consistent with their regulation by other transcription factors. Screening of TRANSFAC matrices indicated enrichment of E2F binding sites in up-regulated secondary targets, as well as of sites for CCAAT box factors, CREB and Forkhead (HNF-3B) transcription factors (Supplementary Tables 1 & 2). In addition, sites for AP1, CREB, C/EBP, CCAAT-box factors, forkhead factors (FOXA, FOXD3, FOXO4, HFH8), E2F, GATA factors, GC-box factors, MAZR, and PAX factors were enriched in down-regulated secondary target genes (Supplementary Tables 1 & 2).

Because secondary target genes are enriched in cell cycle-associated proteins, and of the known role of E2F factors in regulation of cell proliferation, we examined more closely the possible contribution of these factors in the networks of gene regulation by estrogens. Enrichment in E2F sites within a 5 kb window around transcriptional start sites was strong in genes up-regulated in the absence of CHX (Fig. 9A). It was also observed in genes up-regulated in the presence of CHX, but not in the only+CHX group, suggesting that the –and+CHX group of regulated genes was responsible for this over-representation. A similar enrichment in secondary, but not primary target genes was observed when restricting the analysis for each group to genes found to contain E2F binding sites in their promoter in ChIP experiments, although enrichment in the -and+CHX group was not observed. Enrichment in down-regulated secondary genes was not observed when

considering all regulated genes (Fig. 9, Supplementary Table 1), but was detected in the top 100 regulated genes (Supplementary Table 2). Several E2F family members in addition to E2F1 were regulated by E2, including E2F2, 6, 7, and 8 (Fig. 4B), which may explain the prevalence of the E2F signature in regulated secondary genes.

Genes down-regulated in the absence of CHX were enriched in binding sites for NF- κ B (Supplementary Table 1), consistent with the repressive role of estrogens on NF- κ B signaling (15,33) although enrichment of these sites in the absence rather than the presence of CHX was unexpected. Similarly, binding sites for GATA factors were enriched in down-regulated secondary target genes (Supplementary Table 2). Repression of GATA1 transcriptional activity by estrogens has been shown previously (72,73). More surprisingly in view of the reported stimulatory effects of estrogens on Sp1- and AP1- mediated transcription, sites for these transcription factors were also found enriched in down-regulated secondary target genes (Supplementary Tables 1&2). These results suggest that AP1 sites may be involved in mediating gene repression as well as activation (see above for enrichment in up-regulated primary target genes).

Unexpectedly, c-myc binding sites were not found enriched in any of the target gene categories, whether considering all sequences, low complexity sequences or conserved sequences. This is puzzling in view of the observation that induction or repression of MYC expression recapitulates the effects of estrogen and antiestrogens, respectively, on the cell cycle of ER α -positive breast cancer cells (74). However, comparison of E2 target genes with those of Myc target genes (75) and with genes regulated by siRNA-mediated suppression of MYC expression (76) indicates statistically significant enrichment of c-myc target genes in genes regulated by E2 (data not shown). C-myc target genes were distributed in different groups of regulation (up- versus down-, primary versus secondary), which may explain the lack of significant enrichment in a specific group. Poor performance of the TRANSFAC matrices in identifying c-myc binding sites could also play a role in the absence of detected over-representation of these sites in estrogen target genes. It remains

possible also that identification of E2 target genes at earlier time points may reveal a higher enrichment in c-myc binding sites. Of interest, TIAM1, a repressor of c-myc that prevents its apoptotic effects (77), is also a primary estrogen target gene and may limit the duration of the induction of some c-myc target genes.

Together, these data indicate that E2F is a major mediator of propagation of estrogen signaling in MCF7 breast cancer cells for both up- and down-regulated secondary target genes, while AP1 may mediate mostly down-regulation of secondary target genes and Myc may play a more restricted role or be needed at different times.

DISCUSSION

Microarray analyses of estrogen target genes have been performed previously in various ER α positive cell lines. However, only a limited number of genes has been found to be commonly regulated so far (46,70). Although this may be due to intrinsic differences between responses in these cells (70), variability between studies may also take place due to differences in cell culture conditions, microarray platform and selection criteria for significantly regulated genes. In this study, we chose to perform microarray analysis in 0.5% serum, to minimize unstimulated levels of cellular growth; other studies were performed in 5% (45,50) or 10% serum (69,70). Our control was provided by cells treated with vehicle for the same time as with E2, whereas other studies have compared induction by E2 at different times with the absence of treatment at time 0 (50,69,70). We used the HG-U133 2.0 Plus GeneChips, while most other studies used anterior versions of Affymetrix GeneChips (45,69,70). Finally, we used a combination of three parameters, i.e. fold-change, average expression levels and p-value, to rank regulated genes with set minima for each parameter (see Materials and Methods). We identified 2144 significantly regulated genes, of which 17% to 21% overlapped with previous studies (50,70). As we did not observe a higher overlap between sets of regulated genes described in other studies, we

speculate that the low overlap between studies results from a combination of differences in experimental design and data analysis, rather than in the use of cycloheximide specific to our study.

While enrichment in different transcription factors has been reported in E2 targets in different studies (47,69), only one study based on microarray analysis exclusively has reported enrichment in EREs in the promoter sequences (-3 kb to + 0.5 kb) of 85 genes identified as primary targets based on lack of effect of cycloheximide on regulation. Here, we identified 780 genes significantly regulated by E2 in the presence of CHX (544 up-regulated and 236 down-regulated genes). Overlap with genes regulated in the absence of CHX was partial. This can be explained in part by the fact that regulation of several genes found significantly modulated in the presence of CHX was also observed in its absence at earlier times, but became non-significant at 24 h. Genes in this category may include G1 phase-specific genes, as cycloheximide inhibited the increase in S phase induced by E2. In addition, possible inhibition by cycloheximide of negative feed-back regulatory loops may also result in increased gene regulation at 24 h.

Enrichment in EREs was observed only in genes up-regulated in the presence of CHX, including both genes in the absence or presence of CHX and genes regulated in the presence of CHX only. These observations validate the characterization of these genes as primary E2 targets, and explain the lack of reported enrichment of EREs in previous microarray studies that did not discriminate between primary and secondary target genes. Enrichment in EREs in genes up-regulated in the presence of CHX was detectable in a -20/+20 kb window centered on the TSS of regulated genes, consistent with a role of distal EREs in transcriptional regulation.

Regulatory sequences of up-regulated primary target genes often contained several EREs in a 20 kb window around the TSS. For instance, two functional EREs were validated in ChIP experiments in selected primary target genes (CDH26, WISP2, CA12). Of interest,

only two of the EREs validated in ChIP assays were previously reported in genome-wide ChIP-on chip studies (50), likely due to the stringency used in this study to minimize false positive rates. While no correlation between identity with the consensus sequence and apparent strength of binding was observed for individual binding sites, general enrichment in EREs were more pronounced for sequences closer to the consensus sequence, indicating that sequences that are high-affinity binding sites *in vitro* have a higher statistical chance of mediating E2 regulation *in vivo*. The presence of multiple EREs in the vicinity of regulated genes may result in the formation of multiple chromatin loops associating these EREs and the TSS of regulated primary target genes as described in (52).

Surprisingly, genes down-regulated in the presence of CHX were not enriched in EREs. Rather, a statistically significant depletion in EREs was observed. We also observed a lack of enrichment in EREs in ER α -bound chromatin regions located in the vicinity of genes repressed by E2 in ref. (50). This may indicate that the presence of EREs in the flanking regions of genes confers a likelihood of positive rather than negative regulation. However, down-regulation of gene expression via binding of ER α to strategically positioned EREs remains possible, as EREs are still found, although less represented, in the vicinity of repressed primary E2 target genes.

Binding sites for other transcription factors were enriched in primary up- and down-regulated genes. Binding sites enriched in down-regulated primary target genes have not before been associated with E2 action, but include two zinc finger transcription factors participating in corepressor complexes, Evl (78) and Ikaros (IKZF1) (79). On the other hand, binding sites enriched in up-regulated target genes include AP1, which have been previously demonstrated to mediate tethering of ERs. In addition, the ARNT gene encodes a protein that forms a complex with the ligand-bound Ah receptor (AhR). Activated AhR/Arnt was shown to act as a coactivator of estrogen receptor-mediated signaling, while ER α -AHR-ARNT protein-protein interactions mediate estradiol-dependent transrepression of dioxin-inducible gene transcription (80,81). Enrichment of ARNT sites in primary up-

regulated genes suggests that AHR-ARNT-ER α complexes may also mediate up-regulation of some primary estrogen target genes. Finally, EGR factors are, like AP1 components JUN and FOS, immediate early genes activated by a variety of signaling molecules (82,83). It is possible that these factors may contribute to gene regulation via non-genomic activation by estrogen, as suggested before (84).

Sites enriched in primary target genes may mediate regulation either through tethering of ER α or through non-genomic effects. Comparison with sites enriched in ER α -bound chromatin regions may help to discriminate between the two mechanisms, since only transcription factors mediating ER α recruitment via tethering should have sites within ChIP fragments. Enrichment of AP1, Maf, C/EBP, GATA and SOX binding sites in ER α -bound chromatin regions may thus indicate recruitment of ER α to these sites. However, choice of background datasets is more complicated in the case of ChIP-on-chip data, and may affect site enrichment. For instance, enrichment in Egr or Sp1 factors in ER α -associated ChIP fragments was dependent on the choice of the reference background (promoter sequences versus flanking sequences of ChIP regions, data not shown). It should also be noted that, since cooperativity between ER α -ERE complexes and transcription factors bound to the same promoter may result in indirect association of ER α with these transcription factors through chromatin looping, low-levels of enrichment of sites for a variety of transcription factors may be expected in ER α -associated regions.

Several other TFs were found enriched in secondary target genes. Surprisingly, NF- κ B, AP1, GATA and Sp1 sites were all enriched in secondary down-regulated genes. This observation is consistent with the previous report that AP1 sites were present in ER-associated chromatin regions present close to genes repressed by E2 at late time points, this regulation being dependent on the induction of a co-repressor component such as suggested for the NRIP1 ER α co-repressor (50). It will be of interest to determine whether NRIP1 induction is also required for regulation of genes repressed by E2 and containing binding sites for NF- κ B, GATA or Sp1 sites.

One of the most enriched binding sites in up-regulated secondary target genes is that for E2F transcription factors. Enrichment in E2F binding sites is consistent with the well-known role of this family of transcription factors in the control of cell cycle, with known targets including several minichromosome maintenance (MCM) genes, which are essential for cell cycling due to the function of their products as replication-licensing factors (85). Other common E2F and ER targets involved in cell cycle regulation include the CDC6 and CDC25A cell division cycle genes, the PCNA and POLA2 genes, the replication factor C subunit RFC4, the structural maintenance of chromosome 2 (SMC2) and the protein regulator of cytokinesis one (PRC1) genes. Notably, we observed that not only E2F1, but also E2F2, E2F7 and E2F8 transcription factors were up-regulated by E2 with similar kinetics. E2F7 and E2F8 can repress E2F-activated transcription (86,87). E2F6 is also an inhibitory E2F family member (88,89) that can repress c-Myc target genes (90), but its induction was different from that of the other E2F factors, being more rapid and declining in time. The different regulation pattern of this gene may be explained by the fact that it is transcribed in the opposite direction from GREB1, a primary E2 target gene containing three consensus EREs in its flanking sequences. Additional studies will be necessary to investigate the mechanisms of the transcriptional regulation of E2F family members by E2, and their exact individual roles in gene expression and cell cycle regulation by E2.

ACKNOWLEDGEMENTS

This work was supported by grant MOP13-147 from the Canadian Institutes for Health Research to SM. VB has been supported by a CIHR post-doctoral fellowship and JD by an FRSQ studentship. SM and JHW are holders of FRSQ Chercheur-Boursier National awards, and SM holds the CIBC Breast Cancer Research Chair at Université de Montréal.

FIGURE LEGENDS

FIGURE 1: Role of ER α in mediating the effect of estradiol in MCF7 breast cancer cells.

A) MCF7 breast cancer cells were treated every two days with vehicle (0), 17 β -estradiol (E2, 25 nM) or ICI 182,720 (ICI, 100 nM) and protein concentrations were measured after nine days. **B)** Proliferation assays of MCF7 transfected without (0) or with siRNA against ER α , GAPDH or Luciferase. Treatments were initiated 24 h after transfection and protein concentrations were measured 6 days post-transfection. Inset: western blot analysis of ER α and β -actin expression levels in cells transfected with siRNAs directed against ER α or in control cells (T: transfection control). **C)** Cell cycle analysis performed with cells transfected or not with siRNAs for 72 h and stimulated with vehicle or hormone during the last 24 h.

FIGURE 2: Cycloheximide-sensitive and resistant E2 target genes in MCF7 breast cancer cells.

A. Comparative analysis of regulated genes in different large-scale microarray studies. The number of genes in common between this study and those described in refs (50) and (70) or unique to each study is indicated. **B.** Identification of primary and secondary estrogen target genes in MCF7 cells. RNA samples were collected 24 h after treatment of MCF7 cells with vehicle or 17 β -estradiol (25 nM). Cells were pre-treated 1 h before E2 stimulation with cycloheximide (CHX, 10 μ g/ml). Microarray analysis was performed with four replicates

for each condition (see Material & Methods). Numbers of genes significantly regulated (≥ 1.4 fold change between vehicle and E2 treatments, ≥ 5 amplitude of array signals across all conditions and ≤ 0.01 p-value based on moderated t-statistics) are indicated for each category.

FIGURE 3: Expression of several genes induced only in the presence of CHX at 24 h is regulated in its absence at earlier times.

Q-PCR analysis of the effects of E2 (25 nM) and/or CHX (10 $\mu\text{g/ml}$) as a function of time on the expression levels of selected genes identified through microarray analysis. Data is representative of at least two biological experiments. **A)** Pattern of expression of a gene stably induced by estradiol. **B)** Patterns of expression of several genes induced only in the presence of cycloheximide at 24 h in the microarray analysis.

FIGURE 4: Several transcription factors in the E2F family are primary target genes of estradiol in MCF7 breast cancer cells

Q-PCR analysis of the effects of E2 (25 nM) and/or CHX (10 $\mu\text{g/ml}$) as a function of time on FOS (**A**), an immediate early target gene, and of E2F transcription factors (**B**). Induction of most E2Fs is delayed until 6 h or more post-induction and sensitive to cycloheximide, except that of E2F6 whose induction peaks at 2 h and is cycloheximide-resistant. This data is representative of at least two independent experiments.

FIGURE 5: Regulation by estradiol of both up- and down-regulated primary target genes is dependent on ER α .

SiRNA against ER α , GAPDH or Luciferase (LUC) were transiently transfected in MCF7 cells 48 h before stimulation with E2 for 24 h, and levels of expression of ER α (**A**) or of CHX-insensitive (**B**) or sensitive (**C**) up-regulated or down-regulated (**D**) primary target genes were evaluated by quantitative RT-PCR.

FIGURE 6: Estrogen response elements are enriched in the vicinity of primary up-regulated genes.

A) Position weight matrix used for identification of putative EREs in the vicinity of target genes. The frequency of nucleotides at each position of the ERE in the matrix is shown. B) Enrichment in EREs in a 20 kb window (± 10 kb) of the transcriptional start site of estrogen target genes. C) Enrichment in EREs in ER α -bound chromatin regions in ref. (71) found within a 20 kb window of the regulated genes identified in the current study. ChIP regions were standardized to 1 kb on either side of the center of the ChIP region for this analysis.

FIGURE 7: Distribution of EREs in a 50 kb window around the TSS of estrogen target genes.

A) Number of EREs per gene identified at a 75% cut-off rate using the matrix described in Fig. 6 in a 2.5kb sliding window within 50 kb of genomic regions centered around the transcriptional start sites of primary (+ CHX) or total (-CHX) up-regulated genes. B) Number of EREs per gene identified at a 75% cut-off rate in a 2.5kb sliding window within 50 kb genomic regions centered around the transcriptional start sites of primary (+ CHX) or total (-CHX) down-regulated genes. Note that all described alternative TSS were considered in this analysis.

FIGURE 8: Identification of functional EREs in the flanking regions of primary target genes.

Chromatin immunoprecipitation was performed using primer pairs specific for selected estrogen response elements (EREs) identified through bioinformatic screens of the flanking regions of primary E2 target genes. The ChIP assay was performed after 45 min stimulation by vehicle (0), 17 β -estradiol (E2, 25 nM) or ICI 182,720 (ICI, 100 nM) against ER α or non-specific control β -actin. NTC: non-template control for PCR amplification. The number of mismatches (m) with the consensus ERE is indicated for each element along

with distances (bp) to the TSS. These results are representative of two independent experiments.

FIGURE 9: E2F response elements are enriched in secondary up-regulated genes.

A) Relative frequencies in E2F binding sites within conserved regions in a 10 kb window around the transcriptional start sites of regulated genes as a function of the cut-off rates used for matrix-based identification of sites (see Material and Method). Enrichment is observed in up-regulated secondary target genes (not in the only+CHX category). Enrichment in the +CHX group likely reflects overlap with the –CHX group. **B)** Enrichment of E2F sites in a 10 kb window around the transcriptional start sites of genes regulated by E2 in the absence and/or presence of CHX in the current study and reported to contain an E2F1-bound chromatin region in their promoters in ref. (68).

TABLE 1: Top hundred genes regulated by E2 in the presence or absence of cycloheximide.

Genes were ranked based on a combination of three parameters, fold-change between vehicle and E2 treatments, amplitude of array signals across all conditions and p-value based on moderated t-statistics. Genes in bold are those found significantly regulated both in the absence and presence of cycloheximide (CHX). Validated genes are those for which regulation was observed in Q-PCR with at least two independent experiments (see also Fig. 3-5 and supplementary Fig.1; DNS, data not shown).

SUPPLEMENTARY FIGURE 1: Validation of the estrogen regulation of selected genes at 24 h stimulation.

Validation by Q-PCR analysis of the effects of E2 (25 nM) and/or CHX (10 µg/ml) at 24 h on selected genes identified through microarray analysis. The data is representative of at least two independent experiments. A, fold regulation in array analysis; P, fold regulation in quantitative PCR analysis.

SUPPLEMENTARY FIGURE 2: Estradiol-induced proliferation of MCF-7 breast cancer cells is blocked by cycloheximide.

Cell cycle analysis was performed after a 24 hr treatment with vehicle (0) or 17 β -estradiol (E2, 25 nM). Cells were pre-treated with vehicle or CHX (10 μ g/ml) 1 hr before treatment. FACS analysis of propidium iodide stained-MCF7 cells was performed as in Figure 1.

SUPPLEMENTARY TABLE 1: Transcription factor binding sites enriched in a 5 or 10 kb window around the transcriptional start sites of significantly regulated genes.

TRANSFAC matrices were used to screen for enrichment in the different categories of regulated genes with respect to their frequencies in all genes in the human genome within 5 kb (within +/- 2.5 kb from the TSS) or 10 kb (within +/- 5 kb of the TSS). Only matrices corresponding to sites enriched more than 1.5-fold with at least two different cut-offs and significant p-values and Z scores were retained and shown.

SUPPLEMENTARY TABLE 2: Transcription factor binding sites enriched in a 5 or 10 kb window around the transcriptional start sites of the top 100 regulated genes in each category.

Screening with TRANSFAC matrices was performed as in supplementary Table 1, but using the top 100 regulated genes instead of all significantly regulated genes in each category.

REFERENCES

1. Couse, J.F. and Korach, K.S. (1999) Estrogen receptor null mice: what have we learned and where will they lead us? *Endocrine Rev*, **20**, 358-417.
2. McEwen, B.S. and Alves, S.E. (1999) Estrogen actions in the central nervous system. *Endocrine Rev*, **20**, 279-307.
3. Jordan, V.C. (2001) Estrogen, selective estrogen receptor modulation, and coronary heart disease: something or nothing. *J Natl Canc Inst*, **93**, 2-4.
4. Ali, S. and Coombes, R.C. (2002) Endocrine-responsive breast cancer and strategies for combating resistance. *Nat Rev Cancer*, **2**, 101-112.
5. Simpson, E.R. (2003) Sources of estrogen and their importance. *J Steroid Biochem Mol Biol*, **86**, 225-230.
6. Nilsson, S., Makela, S., Treuter, E., Tujague, M., Thomsen, J., Andersson, G., Enmark, E., Pettersson, K., Warner, M. and Gustafsson, J.A. (2001) Mechanisms of estrogen action. *Physiol Rev*, **81**, 1535-1565.
7. Robinson-Rechavi, M., Escriva Garcia, H. and Laudet, V. (2003) The nuclear receptor superfamily. *J Cell Sci*, **116**, 585-586.
8. Kishimoto, M., Fujiki, R., Takezawa, S., Sasaki, Y., Nakamura, T., Yamaoka, K., Kitagawa, H. and Kato, S. (2006) Nuclear receptor mediated gene regulation through chromatin remodeling and histone modifications. *Endocr J*, **53**, 157-172.
9. Metivier, R., Reid, G. and Gannon, F. (2006) Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Rep*, **7**, 161-167.
10. Mader, S., Chambon, P. and White, J.H. (1993) Defining a minimal estrogen receptor DNA binding domain. *Nucleic Acids Res*, **21**, 1125-1132.
11. Schwabe, J.W., Chapman, L., Finch, J.T. and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*, **75**, 567-578.
12. Walker, P., Germond, J.-E., Brown-Luedi, M., Givel, F. and Wahli, W. (1984) Sequence homologies in the region preceding the transcription initiation site of the

- liver estrogen-responsive vitellogenin and apo-VLDLII gene. *Nucleic Acids Res*, **12**, 8611-8626.
13. Klein-Hitpass, L., Ryffel, G.U., Heitlinger, E. and Cato, A.C. (1988) A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. *Nucleic Acids Res*, **16**, 647-663.
 14. Klinge, C.M. (2001) Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Res*, **29**, 2905-2919.
 15. Sanchez, R., Nguyen, D., Rocha, W., White, J.H. and Mader, S. (2002) Diversity in the mechanisms of gene regulation by estrogen receptors. *Bioessays*, **24**, 244-254.
 16. Driscoll, M.D., Sathya, G., Muyan, M., Klinge, C.M., Hilf, R. and Bambara, R.A. (1998) Sequence requirements for estrogen receptor binding to estrogen response elements. *J Biol Chem*, **273**, 29321-29330.
 17. Kulakosky, P.C., McCarty, M.A., Jernigan, S.C., Risinger, K.E. and Klinge, C.M. (2002) Response element sequence modulates estrogen receptor alpha and beta affinity and activity. *J Mol Endocrinol*, **29**, 137-152.
 18. Nguyen, D., Bail, M., Dupont, V., Rocha, W., Deschênes, J., Chteinberg, S. and Mader, S. (2007) Rational design of estrogen receptor mutants with altered DNA binding specificity. *Nucleic Acids Res*, doi: **10.1093/nar/gkm241**.
 19. Bai, Y. and Giguere, V. (2003) Isoform-selective interactions between estrogen receptors and steroid receptor coactivators promoted by estradiol and ErbB-2 signaling in living cells. *Mol Endocrinol*, **17**, 589-599.
 20. Gustafsson, J.A. (2006) ERbeta scientific visions translate to clinical uses. *Climacteric*, **9**, 156-160.
 21. Leo, C. and Chen, J.D. (2000) The SRC family of nuclear receptor coactivators. *Gene*, **245**, 1-11.
 22. Rosenfeld, M.G. and Glass, C.K. (2001) Coregulator codes of transcriptional regulation by nuclear receptors. *J. Biol. Chem.*, **276**, 36865-36868.
 23. Rachez, C. and Freedman, L.P. (2001) Mediator complexes and transcription. *Curr Opin Cell Biol*, **13**, 274-280.

24. Dilworth, F.J. and Chambon, P. (2001) Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene*, **20**, 3047-3054.
25. McDonnell, D.P. and Norris, J.D. (2002) Connections and regulation of the human estrogen receptor. *Science*, **296**, 1642-1644.
26. McKenna, N.J. and O'Malley, B.W. (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell*, **108**, 465-474.
27. Belandia, B. and Parker, M.G. (2003) Nuclear receptors: a rendezvous for chromatin remodeling factors. *Cell*, **114**, 277-280.
28. Forster, C., Makela, S., Warri, A., Kietz, S., Becker, D., Hultenby, K., Warner, M. and Gustafsson, J.A. (2002) Involvement of estrogen receptor beta in terminal differentiation of mammary gland epithelium. *Proc Natl Acad Sci U S A*, **99**, 15578-15583.
29. Leygue, E., Dotzlaw, H., Watson, P.H. and Murphy, L.C. (1998) Altered estrogen receptor alpha and beta messenger RNA expression during human breast tumorigenesis. *Cancer Res*, **58**, 3197-3201.
30. Roger, P., Sahla, M.E., Makela, S., Gustafsson, J.A., Baldet, P. and Rochefort, H. (2001) Decreased expression of estrogen receptor beta protein in proliferative preinvasive mammary tumors. *Cancer Res*, **61**, 2537-2541.
31. Kushner, P.J., Agard, D.A., Greene, G.L., Scanlan, T.S., Shiau, A.K., Uht, R.M. and Webb, P. (2000) Estrogen receptor pathways to AP-1. *J Steroid Biochem Mol Biol*, **74**, 311-317.
32. Safe, S. (2001) Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions. *Vitam. Horm.*, **62**, 231-252.
33. Kalaitzidis, D. and Gilmore, T.D. (2005) Transcription factor cross-talk: the estrogen receptor and NF-kappaB. *Trends Endocrinol Metab*, **16**, 46-52.
34. Migliaccio, A., Di Domenico, M., Castoria, G., de Falco, A., Bontempo, P., Nola, E. and Auricchio, F. (1996) Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells. *Embo J*, **15**, 1292-1300.

35. Sun, M., Paciga, J.E., Feldman, R.I., Yuan, Z., Coppola, D., Lu, Y.Y., Shelley, S.A., Nicosia, S.V. and Cheng, J.Q. (2001) Phosphatidylinositol-3-OH Kinase (PI3K)/AKT2, activated in breast cancer, regulates and is induced by estrogen receptor alpha (ERalpha) via interaction between ERalpha and PI3K. *Cancer Res*, **61**, 5985-5991.
36. Pedram, A., Razandi, M., Aitkenhead, M., Hughes, C.C. and Levin, E.R. (2002) Integration of the non-genomic and genomic actions of estrogen. Membrane-initiated signaling by steroid to transcription and cell biology. *J Biol Chem*, **277**, 50768-50775.
37. Simoncini, T., Fornari, L., Mannella, P., Varone, G., Caruso, A., Liao, J.K. and Genazzani, A.R. (2002) Novel non-transcriptional mechanisms for estrogen receptor signaling in the cardiovascular system. Interaction of estrogen receptor alpha with phosphatidylinositol 3-OH kinase. *Steroids*, **67**, 935-939.
38. Pedram, A., Razandi, M. and Levin, E.R. (2006) Nature of functional estrogen receptors at the plasma membrane. *Mol Endocrinol*, **20**, 1996-2009.
39. Filardo, E.J., Quinn, J.A., Bland, K.I. and Frackelton, A.R., Jr. (2000) Estrogen-induced activation of Erk-1 and Erk-2 requires the G protein-coupled receptor homolog, GPR30, and occurs via trans-activation of the epidermal growth factor receptor through release of HB-EGF. *Mol Endocrinol*, **14**, 1649-1660.
40. Maggiolini, M., Vivacqua, A., Fasanella, G., Recchia, A.G., Sisci, D., Pezzi, V., Montanaro, D., Musti, A.M., Picard, D. and Ando, S. (2004) The G protein-coupled receptor GPR30 mediates c-fos up-regulation by 17beta-estradiol and phytoestrogens in breast cancer cells. *J Biol Chem*, **279**, 27008-27016.
41. Revankar, C.M., Cimino, D.F., Sklar, L.A., Arterburn, J.B. and Prossnitz, E.R. (2005) A transmembrane intracellular estrogen receptor mediates rapid cell signaling. *Science*, **307**, 1625-1630.
42. Thomas, P., Pang, Y., Filardo, E.J. and Dong, J. (2005) Identity of an estrogen membrane receptor coupled to a G protein in human breast cancer cells. *Endocrinology*, **146**, 624-632.

43. Funakoshi, T., Yanai, A., Shinoda, K., Kawano, M.M. and Mizukami, Y. (2006) G protein-coupled receptor 30 is an estrogen receptor in the plasma membrane. *Biochem Biophys Res Commun*, **346**, 904-910.
44. Soulez, M. and Parker, M.G. (2001) Identification of novel oestrogen receptor target genes in human ZR75-1 breast cancer cells by expression profiling. *J. Mol. Endocrinol.*, **27**, 259-274.
45. Frasor, J., Danes, J.M., Komm, B., Chang, K.C., Lyttle, C.R. and Katzenellenbogen, B.S. (2003) Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology*, **144**, 4562-4574.
46. Lin, C.Y., Strom, A., Vega, V.B., Kong, S.L., Yeo, A.L., Thomsen, J.S., Chan, W.C., Doray, B., Bangarusamy, D.K., Ramasamy, A. *et al.* (2004) Discovery of estrogen receptor alpha target genes and response elements in breast tumor cells. *Genome Biol*, **5**, R66.
47. Cicatiello, L., Scafoglio, C., Altucci, L., Cancemi, M., Natoli, G., Facchiano, A., Iazzetti, G., Calogero, R., Biglia, N., De Bortoli, M. *et al.* (2004) A genomic view of estrogen actions in human breast cancer cells by expression profiling of the hormone-responsive transcriptome. *J Mol Endocrinol*, **32**, 719-775.
48. Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R. *et al.* (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell*, **122**, 33-43.
49. Vega, V.B., Lin, C.Y., Lai, K.S., Kong, S.L., Xie, M., Su, X., Teh, H.F., Thomsen, J.S., Yeo, A.L., Sung, W.K. *et al.* (2006) Multiplatform genome-wide identification and modeling of functional human estrogen receptor binding sites. *Genome Biol*, **7**, R82.
50. Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F. *et al.* (2006) Genome-wide analysis of estrogen receptor binding sites. *Nat Genet*, **38**, 1289-1297.

51. Bourdeau, V., Deschenes, J., Metivier, R., Nagai, Y., Nguyen, D., Bretschneider, N., Gannon, F., White, J.H. and Mader, S. (2004) Genome-wide identification of high-affinity estrogen response elements in human and mouse. *Mol Endocrinol*, **18**, 1411-1427.
52. Deschenes, J., Bourdeau, V., White, J.H. and Mader, S. (2007) Regulation of GREB1 transcription by estrogen receptor alpha through a multipartite enhancer spread over 20 KB of upstream flanking sequences. *J Biol Chem*.
53. Rocha, W., Sanchez, R., Deschenes, J., Auger, A., Hebert, E., White, J.H. and Mader, S. (2005) Opposite effects of histone deacetylase inhibitors on glucocorticoid and estrogen signaling in human endometrial Ishikawa cells. *Mol Pharmacol*, **68**, 1852-1862.
54. Dayan, G., Lupien, M., Auger, A., Anghel, S.I., Rocha, W., Croisette, S., Katzenellenbogen, J.A. and Mader, S. (2006) Tamoxifen and raloxifene differ in their functional interactions with aspartate 351 of estrogen receptor alpha. *Mol Pharmacol*, **70**, 579-588.
55. Wettenhall, J.M., Simpson, K.M., Satterley, K. and Smyth, G.K. (2006) affyImGUI: a graphical user interface for linear modeling of single channel microarray data. *Bioinformatics*, **22**, 897-899.
56. Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*, **31**, e15.
57. Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J. *et al.* (2003) The UCSC Genome Browser Database. *Nucleic Acids Res*, **31**, 51-54.
58. Pruitt, K.D., Tatusova, T. and Maglott, D.R. (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*, **33**, D501-504.
59. Wasserman, W.W. and Sandelin, A. (2004) Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet*, **5**, 276-287.

60. Wang, T.T., Tavera-Mendoza, L.E., Laperriere, D., Libby, E., Macleod, N.B., Nagai, Y., Bourdeau, V., Konstorum, A., Lallement, B., Zhang, R. *et al.* (2005) Large-scale in silico and microarray-based identification of direct 1,25-dihydroxyvitamin D3 target genes. *Mol Endocrinol*.
61. Laperriere, D., Wang, T.T., White, J.H. and Mader, S. (2007) Widespread Alu repeat-driven expansion of consensus DR2 retinoic acid response elements during primate evolution. *BMC Genomics*, **8**, 23.
62. Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K. *et al.* (2006) TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res*, **34**, D108-110.
63. Jin, V.X., Sun, H., Pohar, T.T., Liyanarachchi, S., Palaniswamy, S.K., Huang, T.H. and Davuluri, R.V. (2005) ERTargetDB: an integral information resource of transcription regulation of estrogen receptor target genes. *J Mol Endocrinol*, **35**, 225-230.
64. O'Lone, R., Frith, M.C., Karlsson, E.K. and Hansen, U. (2004) Genomic targets of nuclear estrogen receptors. *Mol Endocrinol*, **18**, 1859-1875.
65. Balmer, J.E. and Blomhoff, R. (2005) A robust characterization of retinoic acid response elements based on a comparison of sites in three species. *J Steroid Biochem Mol Biol*, **96**, 347-354.
66. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*, **15**, 1034-1050.
67. Ho Sui, S.J., Mortimer, J.R., Arenillas, D.J., Brumm, J., Walsh, C.J., Kennedy, B.P. and Wasserman, W.W. (2005) oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic Acids Res*, **33**, 3154-3164.

68. Jin, V.X., Rabinovich, A., Squazzo, S.L., Green, R. and Farnham, P.J. (2006) A computational genomics approach to identify cis-regulatory modules from chromatin immunoprecipitation microarray data--a case study using E2F1. *Genome Res*, **16**, 1585-1595.
69. Creighton, C.J., Cordero, K.E., Larios, J.M., Miller, R.S., Johnson, M.D., Chinnaiyan, A.M., Lippman, M.E. and Rae, J.M. (2006) Genes regulated by estrogen in breast tumor cells in vitro are similarly regulated in vivo in tumor xenografts and human breast tumors. *Genome Biol*, **7**, R28.
70. Rae, J.M., Johnson, M.D., Scheys, J.O., Cordero, K.E., Larios, J.M. and Lippman, M.E. (2005) GREB 1 is a critical regulator of hormone dependent breast cancer growth. *Breast Cancer Res Treat*, **92**, 141-149.
71. Carroll, J.S. and Brown, M. (2006) Estrogen receptor target gene: an evolving concept. *Mol Endocrinol*, **20**, 1707-1714.
72. Blobel, G.A., Sieff, C.A. and Orkin, S.H. (1995) Ligand-dependent repression of the erythroid transcription factor GATA-1 by the estrogen receptor. *Mol Cell Biol*, **15**, 3147-3153.
73. Holth, L.T., Sun, J.M., Coutts, A.S., Murphy, L.C. and Davie, J.R. (1997) Estrogen receptor diminishes DNA-binding activities of chicken GATA-1 and CACCC-binding proteins. *DNA Cell Biol*, **16**, 1477-1482.
74. Doisneau-Sixou, S.F., Sergio, C.M., Carroll, J.S., Hui, R., Musgrove, E.A. and Sutherland, R.L. (2003) Estrogen and antiestrogen regulation of cell cycle progression in breast cancer cells. *Endocr Relat Cancer*, **10**, 179-186.
75. Zhao, F., Xuan, Z., Liu, L. and Zhang, M.Q. (2005) TRED: a Transcriptional Regulatory Element Database and a platform for in silico gene regulation studies. *Nucleic Acids Res*, **33**, D103-107.
76. Cappellen, D., Schlange, T., Bauer, M., Maurer, F. and Hynes, N.E. (2007) Novel c-MYC target genes mediate differential effects on cell proliferation and migration. *EMBO Rep*, **8**, 70-76.

77. Otsuki, Y., Tanaka, M., Kamo, T., Kitanaka, C., Kuchino, Y. and Sugimura, H. (2003) Guanine nucleotide exchange factor, Tiam1, directly binds to c-Myc and interferes with c-Myc-mediated apoptosis in rat-1 fibroblasts. *J Biol Chem*, **278**, 5132-5140.
78. Hirai, H., Izutsu, K., Kurokawa, M. and Mitani, K. (2001) Oncogenic mechanisms of Evi-1 protein. *Cancer Chemother Pharmacol*, **48 Suppl 1**, S35-40.
79. Koipally, J., Renold, A., Kim, J. and Georgopoulos, K. (1999) Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. *Embo J*, **18**, 3090-3100.
80. Brunnberg, S., Pettersson, K., Rydin, E., Matthews, J., Hanberg, A. and Pongratz, I. (2003) The basic helix-loop-helix-PAS protein ARNT functions as a potent coactivator of estrogen receptor-dependent transcription. *Proc Natl Acad Sci U S A*, **100**, 6517-6522.
81. Ohtake, F., Takeyama, K., Matsumoto, T., Kitagawa, H., Yamamoto, Y., Nohara, K., Tohyama, C., Krust, A., Mimura, J., Chambon, P. *et al.* (2003) Modulation of oestrogen receptor signalling by association with the activated dioxin receptor. *Nature*, **423**, 545-550.
82. Sukhatme, V.P. (1990) Early transcriptional events in cell growth: the Egr family. *J Am Soc Nephrol*, **1**, 859-866.
83. Khachigian, L.M. and Collins, T. (1998) Early growth response factor 1: a pleiotropic mediator of inducible gene expression. *J Mol Med*, **76**, 613-616.
84. Chen, C.C., Lee, W.R. and Safe, S. (2004) Egr-1 is activated by 17beta-estradiol in MCF-7 cells by mitogen-activated protein kinase-dependent phosphorylation of ELK-1. *J Cell Biochem*, **93**, 1063-1074.
85. Kato, K., Toki, T., Shimizu, M., Shiozawa, T., Fujii, S., Nikaido, T. and Konishi, I. (2003) Expression of replication-licensing factors MCM2 and MCM3 in normal, hyperplastic, and carcinomatous endometrium: correlation with expression of Ki-67 and estrogen and progesterone receptors. *Int J Gynecol Pathol*, **22**, 334-340.

86. Di Stefano, L., Jensen, M.R. and Helin, K. (2003) E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. *Embo J*, **22**, 6289-6298.
87. Christensen, J., Cloos, P., Toftegaard, U., Klinkenberg, D., Bracken, A.P., Trinh, E., Heeran, M., Di Stefano, L. and Helin, K. (2005) Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. *Nucleic Acids Res*, **33**, 5458-5470.
88. Trimarchi, J.M., Fairchild, B., Verona, R., Moberg, K., Andon, N. and Lees, J.A. (1998) E2F-6, a member of the E2F family that can behave as a transcriptional repressor. *Proc Natl Acad Sci U S A*, **95**, 2850-2855.
89. Cartwright, P., Muller, H., Wagener, C., Holm, K. and Helin, K. (1998) E2F-6: a novel member of the E2F family is an inhibitor of E2F-dependent transcription. *Oncogene*, **17**, 611-623.
90. Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M. and Nakatani, Y. (2002) A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science*, **296**, 1132-1136.

Figure 1

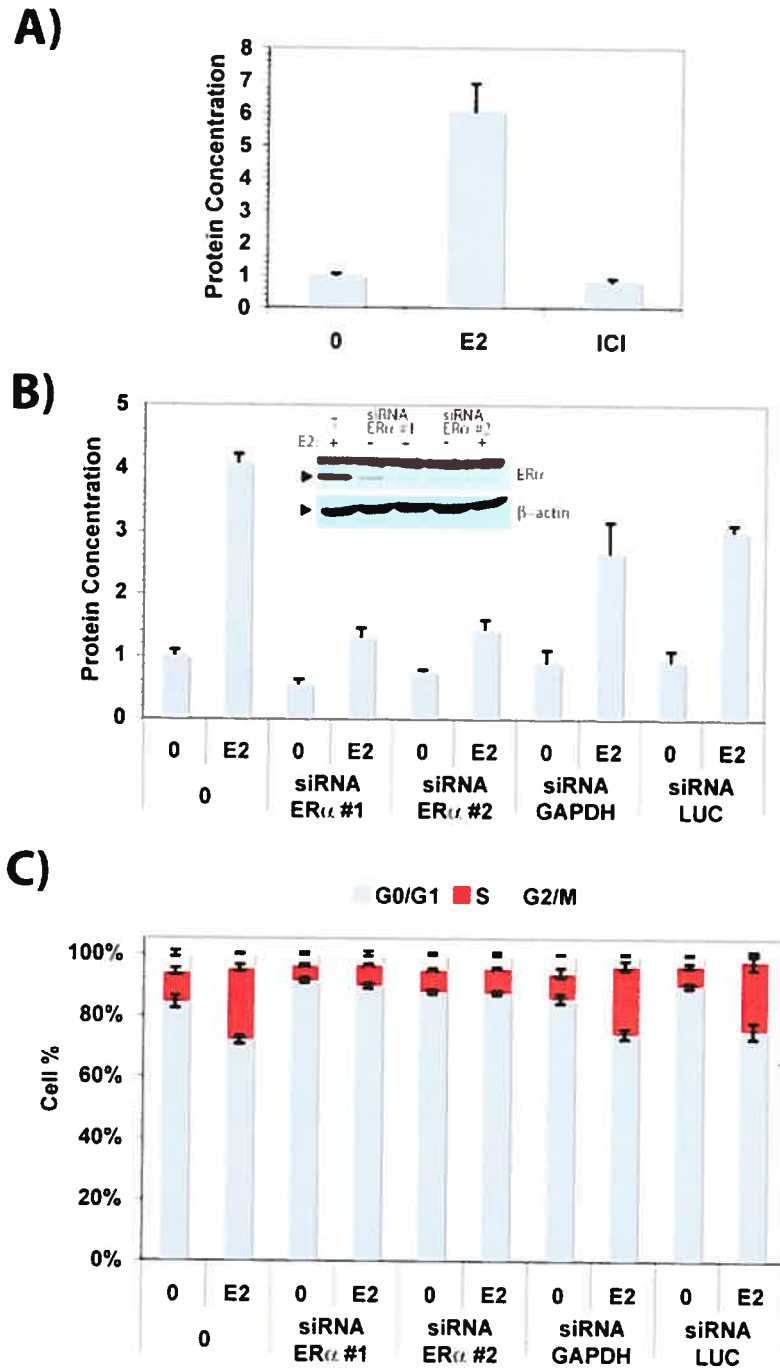


Figure 2

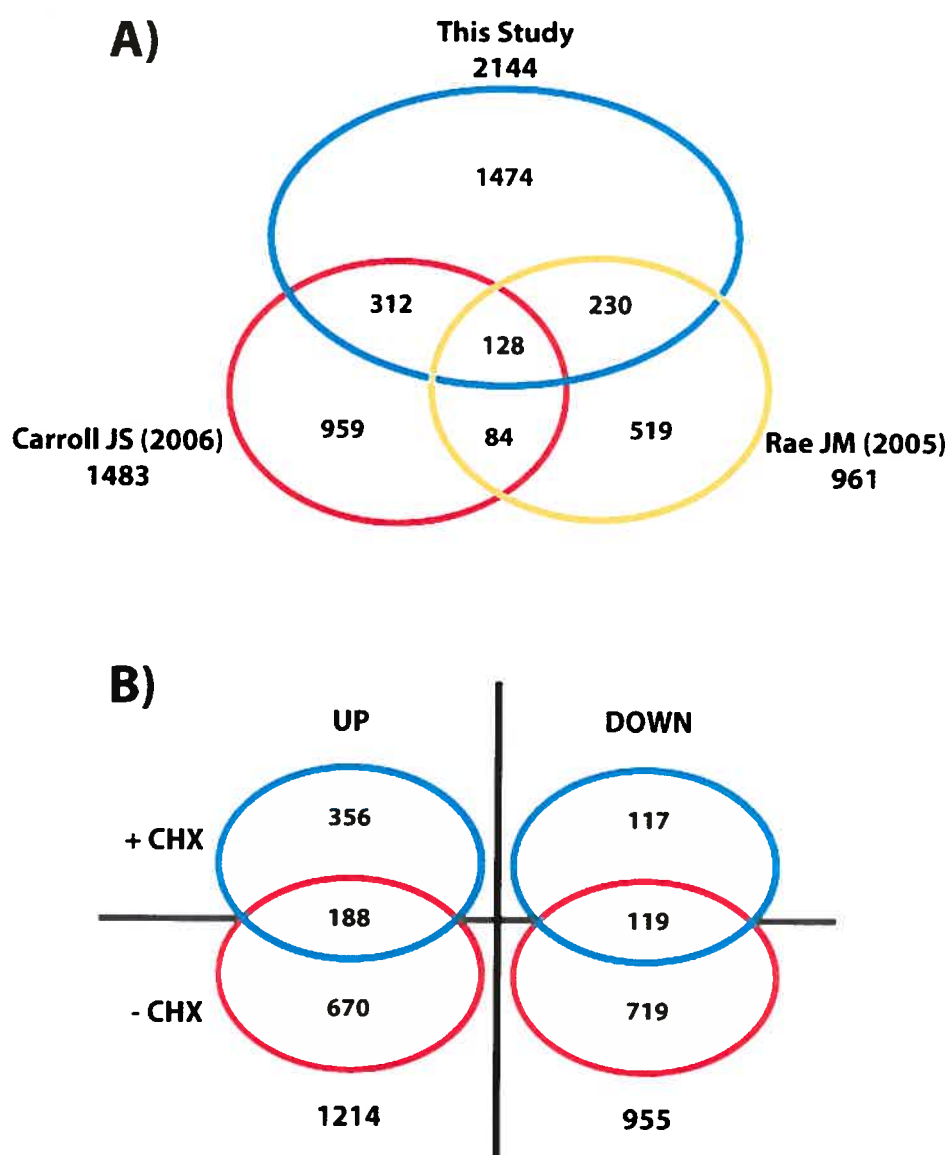


Figure 3

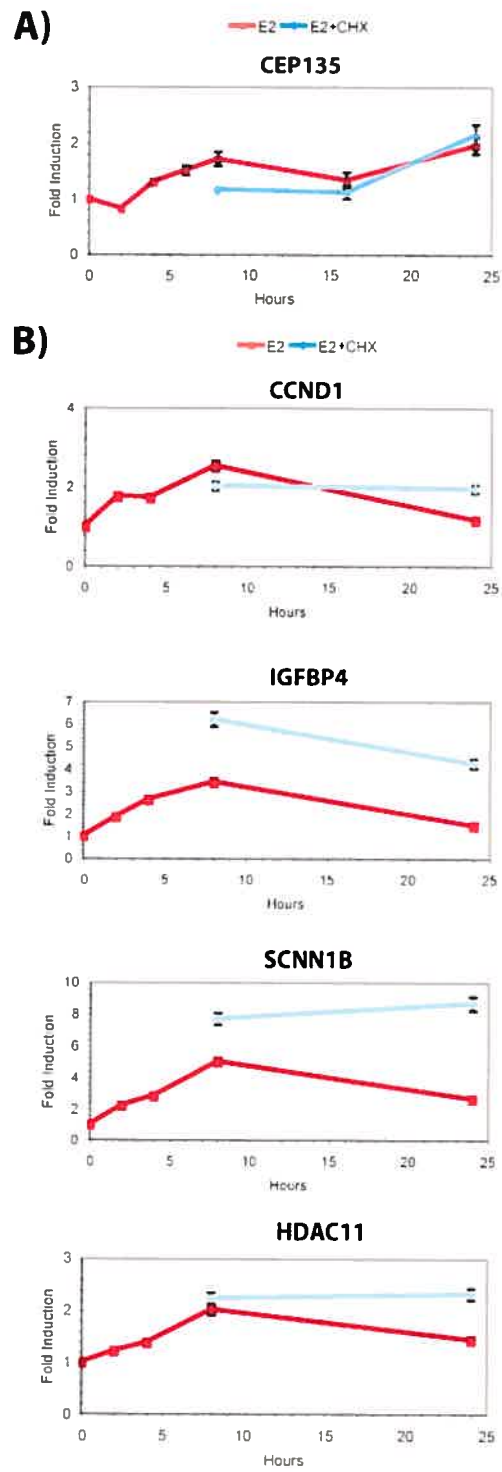


Figure 4

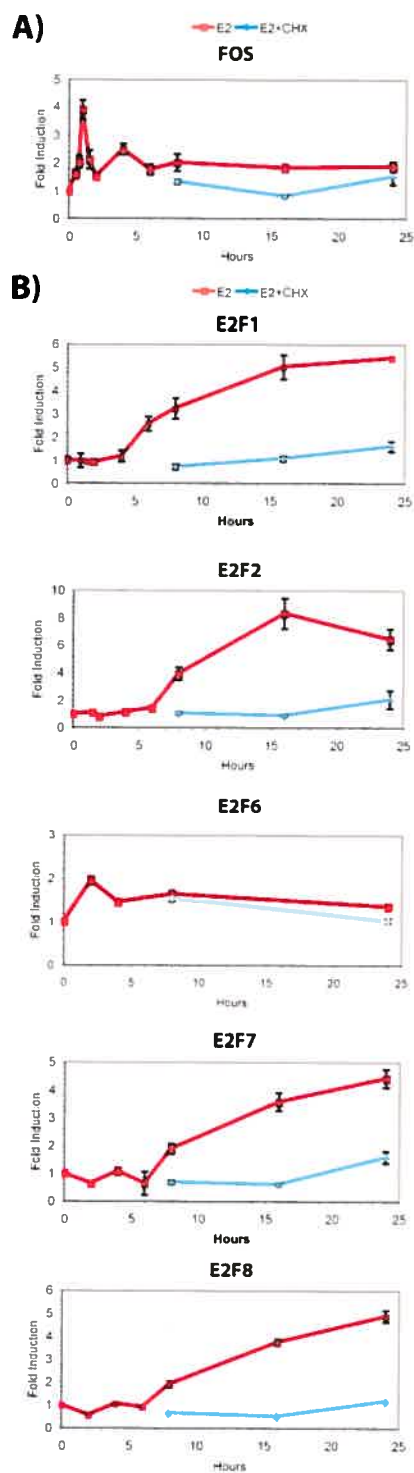


Figure 5

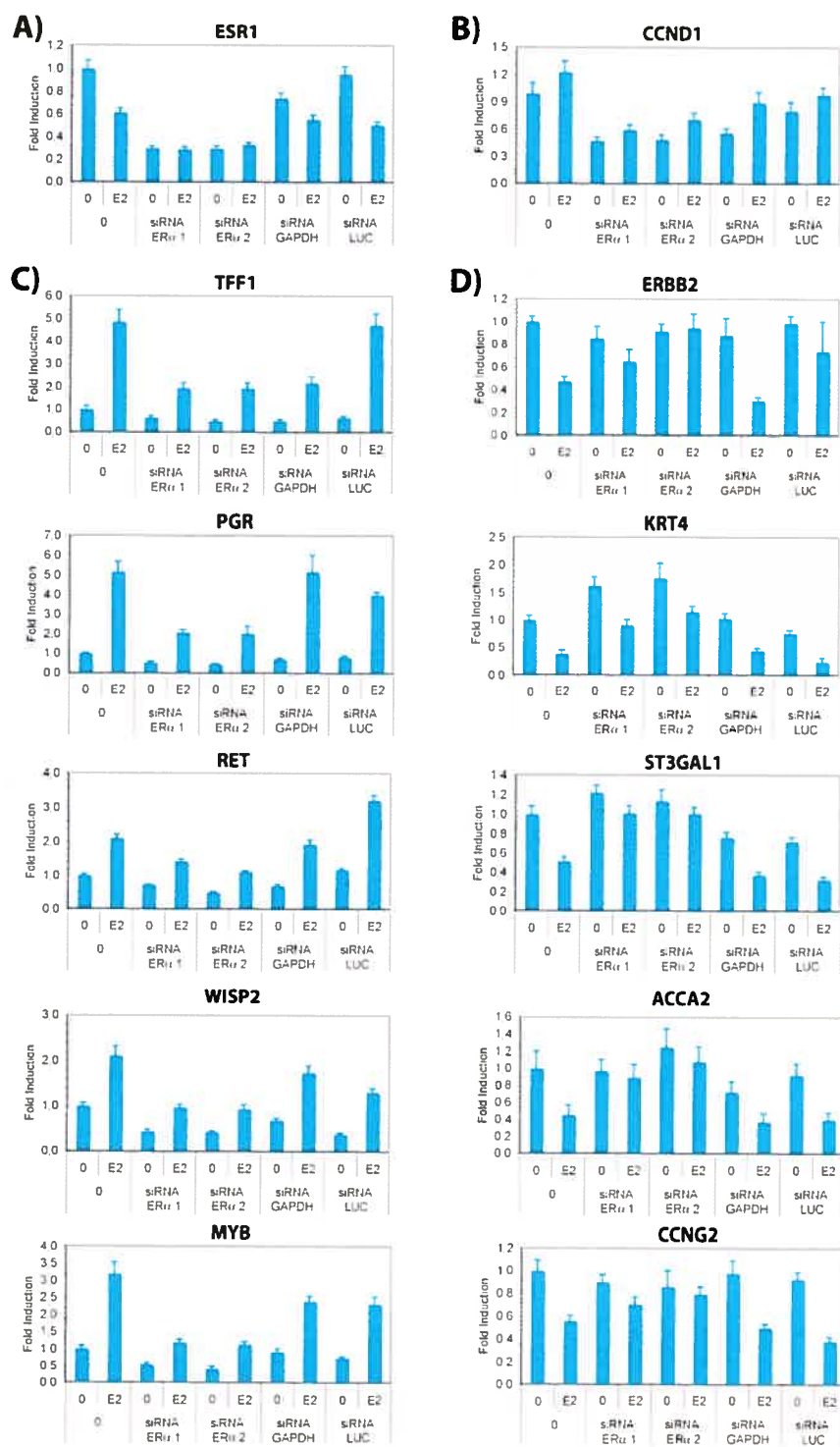


Figure 6

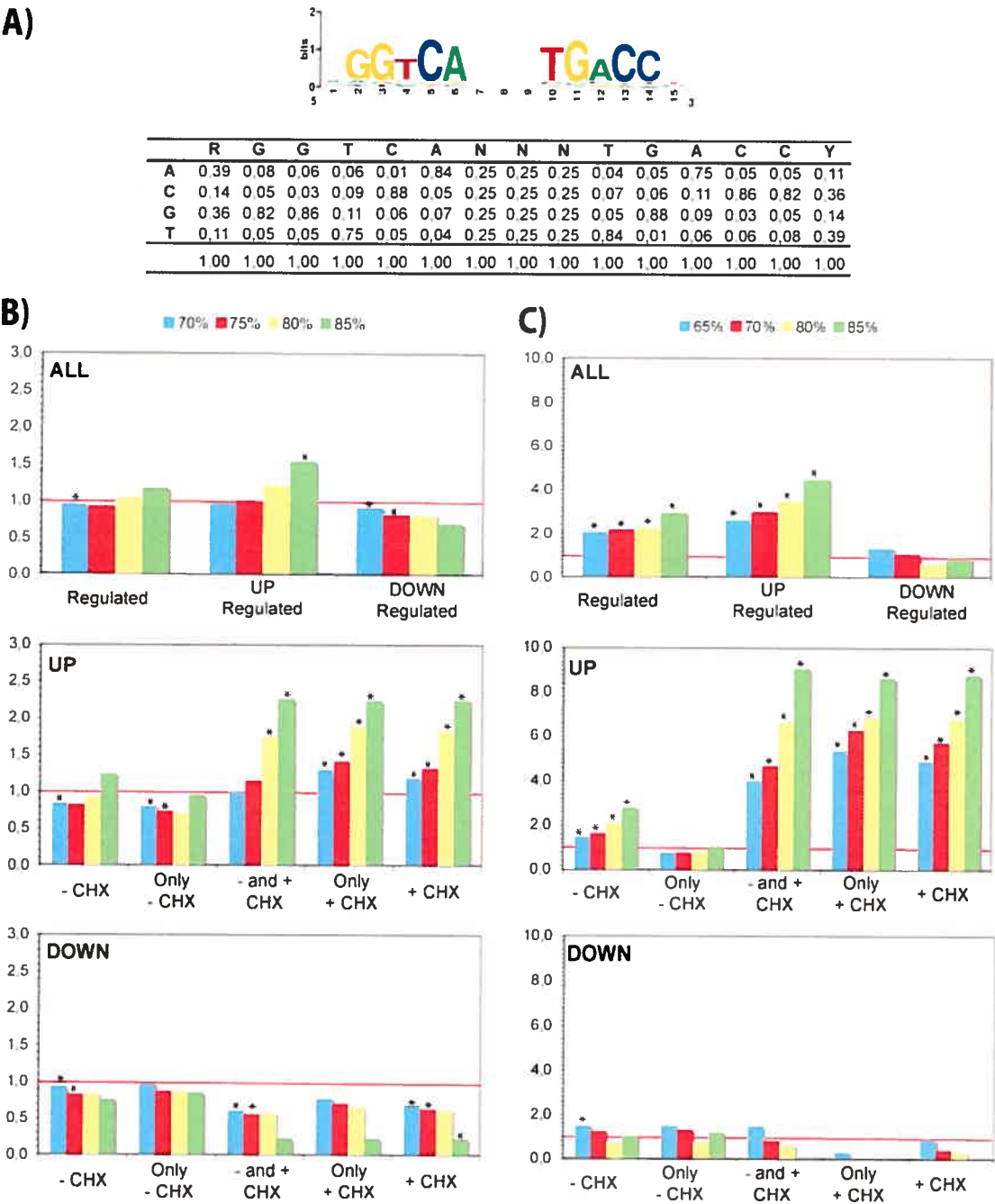


Figure 7

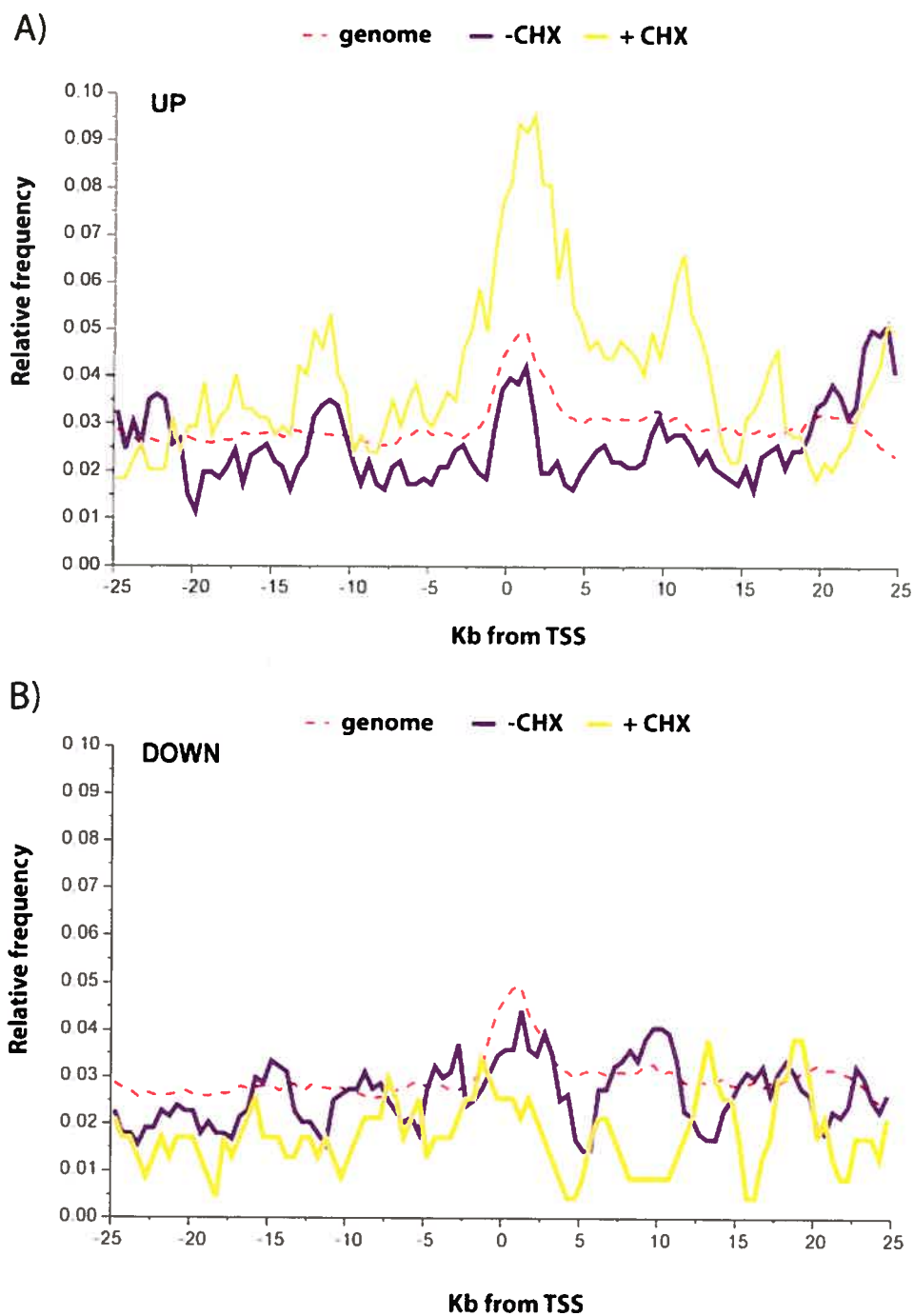


Figure 8

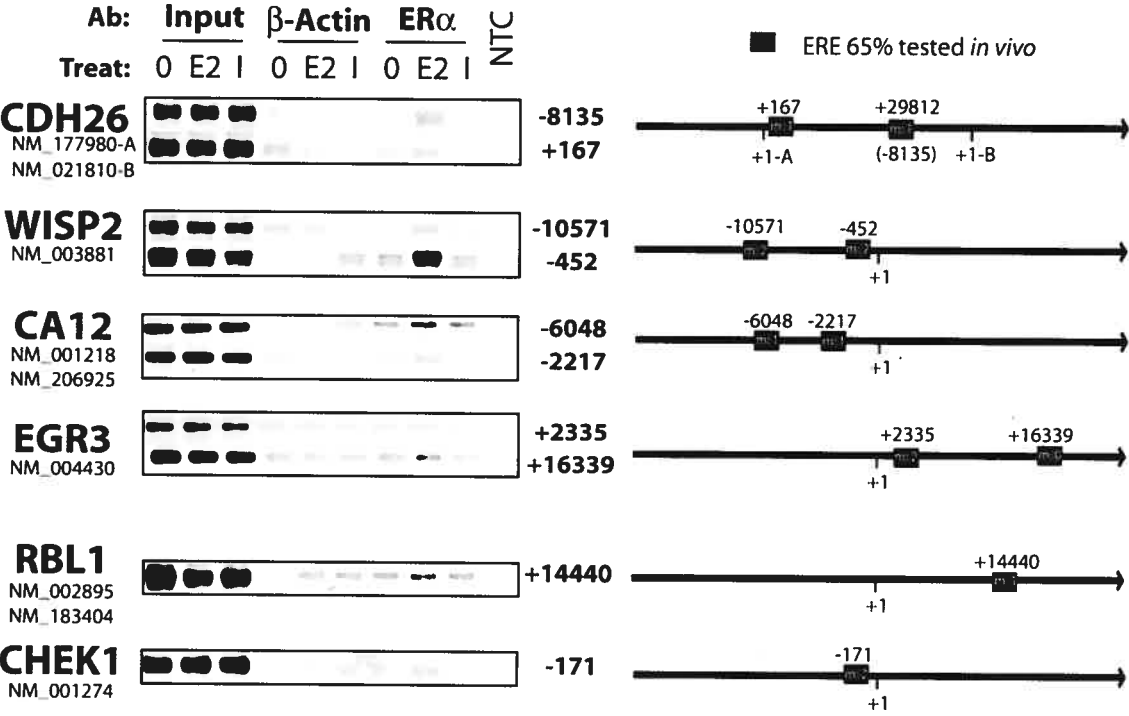


Figure 9

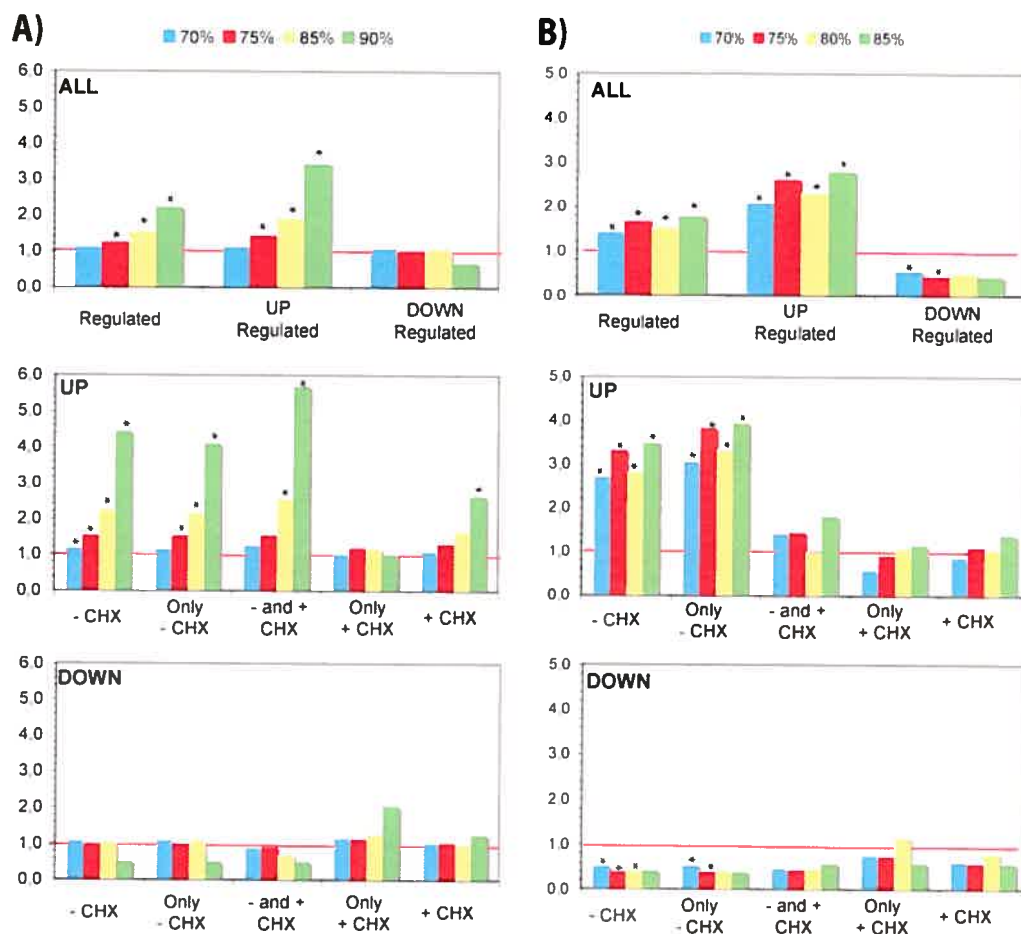


Table 1

Without CHX						With CHX					
RANK	Symbol	Fold	A	P-Value	Validated	RANK	Symbol	Fold	A	P-Value	Validated
1	MYBL1	13.85	7.23	4.07E-10	yes	1	KRT13	15.90	6.81	3.46E-13	-
2	FLJ30058	14.85	5.66	5.16E-12	-	2	LOC283551	15.04	5.81	3.26E-13	-
3	PKIB	7.09	10.22	5.01E-13	yes	3	FLJ18847	11.11	5.12	1.46E-16	-
4	SYTL4	5.79	7.86	4.39E-10	-	4	CYP26B1	7.46	7.05	5.91E-12	yes (DNS)
5	SGK3	5.99	6.93	2.40E-08	-	5	SGK3	6.89	6.93	1.86E-08	yes (DNS)
6	UGT2B15	6.19	6.45	3.14E-09	-	6	FLJ30058	7.60	5.66	4.93E-11	-
7	PCP4	5.03	7.28	1.14E-11	-	7	SLC26A2	5.40	7.05	1.22E-11	yes
8	TM6RSS3	5.42	6.72	9.36E-12	-	8	MYBL1	5.27	7.23	3.45E-07	yes
9	MCM10	4.96	6.97	8.87E-11	-	9	RET	4.68	8.12	5.30E-12	yes
10	RRM2	3.44	10.03	9.07E-12	yes	10	C2orf46	6.10	5.85	8.37E-12	-
11	MLF1P	3.89	8.76	1.56E-11	-	11	RBP7	4.53	7.75	5.91E-12	yes
12	ATAD2	3.47	9.45	6.83E-10	-	12	RBM24	4.95	7.02	7.32E-12	-
13	DCC1	4.78	6.54	1.34E-10	-	13	TMEM64	3.26	10.13	3.50E-10	-
14	UHRF1	3.49	9.04	2.63E-11	-	14	CA12	3.43	9.42	5.66E-13	yes
15	TK1	3.58	8.80	4.97E-11	-	15	XBP1	3.13	10.27	4.50E-11	yes
16	IL17RB	4.24	7.38	1.87E-11	-	16	KCNK5	4.33	7.32	5.30E-12	-
17	GREB1	3.14	9.85	5.09E-11	yes	17	DEPDC6	3.75	8.25	8.37E-12	-
18	PBK	3.93	7.85	8.91E-10	-	18	IL17RB	4.08	7.33	2.12E-11	-
19	BRIP1	3.08	9.84	2.39E-07	-	19	OACT1	4.68	6.32	8.37E-12	-
20	TYMS	3.28	9.18	5.91E-11	-	20	CDN26	4.96	5.82	1.26E-07	yes
21	DTL	4.49	6.66	4.39E-10	-	21	CALCR	2.82	9.97	3.18E-11	yes
22	EXO1	4.54	6.58	9.07E-12	-	22	CELSR2	3.22	8.66	1.41E-12	yes (DNS)
23	LOC221981	2.89	10.32	3.88E-10	-	23	KRT4	3.91	7.01	1.53E-08	yes
24	ANLN	3.34	8.90	8.05E-09	-	24	PAPSS2	2.60	10.44	1.03E-10	-
25	CNE2	3.69	7.89	1.95E-09	-	25	DSCAM	4.12	6.57	1.05E-10	yes
26	CDC2	3.10	9.34	1.98E-10	-	26	CARD10	4.04	6.59	1.59E-11	yes (DNS)
27	AREG	3.32	8.69	2.35E-11	yes	27	TPD52L1	2.55	10.14	5.69E-11	-
28	RLN2	3.38	8.36	1.11E-09	-	28	CXCL12	3.62	6.99	8.64E-10	-
29	MOG	3.79	7.27	4.63E-11	-	29	LRIG1	3.82	6.54	5.91E-12	yes
30	CXCL12	3.97	6.99	2.28E-10	-	30	RLN2	2.97	8.36	9.91E-09	-
31	CHBK1	4.18	6.61	7.62E-09	Yes (DNS)	31	ANKK4	3.77	5.99	3.18E-11	-
32	TOP2A	2.97	9.24	1.60E-09	-	32	C14orf139	4.71	5.08	5.30E-12	-
33	KIAA0101	2.63	10.41	1.38E-10	-	33	C22orf19	2.90	8.23	2.37E-11	-
34	RERG	3.33	8.15	9.07E-12	-	34	IP13-360B22	2.88	8.21	4.39E-11	-
35	CYP26B1	3.85	7.05	6.70E-10	yes	35	PDX	3.59	6.53	4.33E-09	-
36	FSHPRH1	4.41	6.15	1.30E-10	-	36	TUBA3	2.18	10.62	3.25E-09	-
37	Pfz2	3.31	8.19	1.24E-10	-	37	MYB	2.62	8.66	5.91E-12	yes
38	BRCA1	4.15	6.49	2.17E-10	yes	38	C2orf23	3.04	7.43	5.91E-12	-
39	CDKN3	2.90	9.28	1.33E-09	-	39	UGT2B15	3.47	6.45	1.20E-06	-
40	SFXN2	3.15	8.51	1.87E-11	-	40	STC2	2.38	9.31	2.50E-11	-
41	FLJ10719	3.69	7.25	1.60E-09	-	41	P4HA2	2.62	8.29	5.59E-11	-
42	CDCA5L	3.82	6.86	1.14E-11	yes	42	FLJ33718	2.73	7.92	1.03E-10	-
43	PCNA	2.61	10.01	2.54E-10	-	43	SMOX	3.20	6.72	2.81E-11	-
44	CDH2	2.68	9.71	2.48E-11	-	44	GPM1	2.52	8.13	8.27E-11	-
45	HEY2	3.46	7.51	2.48E-11	-	45	GREB1	2.15	9.85	1.62E-08	yes
46	C10orf3	3.47	7.47	4.06E-08	-	46	ITGA2	3.52	5.99	9.34E-10	-
47	G30	2.74	9.40	1.21E-11	-	47	FKBP4	1.90	10.90	2.01E-09	-
48	GMNN	2.99	8.60	4.97E-11	-	48	PKIB	2.00	10.22	1.22E-07	yes
49	DEPDC6	3.05	8.35	6.81E-11	-	49	RAB27B	2.33	8.70	1.78E-08	yes
50	CDG5	4.21	6.05	3.51E-09	-	50	HAA1324	2.10	9.64	1.79E-10	-
51	KLK11	2.74	9.28	2.48E-11	-	51	PCP4	2.78	7.28	4.14E-09	-
52	FKSG14	4.34	5.82	1.10E-08	-	52	UGCG	2.54	7.56	2.44E-08	-
53	FHL1	2.93	8.63	4.39E-10	-	53	RHOD	2.03	9.83	1.11E-09	-
54	MAD2L1	3.14	8.04	6.55E-09	-	54	ANKA9	2.00	9.92	2.47E-09	-
55	PRDM1	3.22	7.84	4.63E-11	-	55	WFS1	2.43	8.16	7.02E-10	-
56	GLA	2.78	9.06	1.14E-11	-	56	SLC29A1	3.22	6.14	2.85E-10	-
57	CDT1	2.98	8.42	1.49E-11	-	57	SOX9	2.47	8.02	1.70E-09	-
58	MYBL2	3.15	7.89	9.57E-11	yes	58	PLEKHA7	2.98	6.87	8.27E-11	-
59	BCMP11	3.27	7.57	2.37E-09	-	59	SH2BP1	2.33	8.45	9.51E-09	yes
60	HCAP-G	3.54	6.99	8.72E-10	-	60	WSP2	2.04	9.65	8.39E-10	yes
61	RAD51AP1	3.97	6.20	3.32E-08	-	61	SLC9A3R1	1.76	11.14	7.70E-09	-
62	CDH26	4.28	5.82	2.32E-07	yes	62	RERG	2.41	8.15	2.90E-10	Yes (DNS)
63	ZNF387	3.30	7.33	1.38E-10	-	63	IGFBP4	1.97	9.88	7.97E-09	yes
64	CDCA5	3.06	7.91	4.85E-11	-	64	CD33L3	2.82	6.89	4.17E-10	-
65	ASF1B	2.97	8.10	1.38E-11	-	65	RNF150	3.57	5.41	3.97E-10	-
66	UBE2C	2.38	10.08	4.02E-10	-	66	TGM2	2.97	6.48	1.85E-08	-
67	UBE2T	2.63	9.11	2.48E-11	-	67	SOX2	2.21	8.69	1.83E-08	-
68	KNTC2	3.32	7.16	2.71E-07	-	68	SCARB1	2.27	8.44	1.86E-10	-
69	MELK	3.07	7.70	1.19E-10	-	69	ABHD2	2.44	7.85	1.10E-09	-
70	SPBC25	4.36	5.42	3.51E-09	-	70	BLVRB	1.97	9.71	8.39E-10	-
71	LOC440887	2.90	8.12	3.99E-09	-	71	IDH1	2.80	9.54	2.89E-10	-
72	E2F8	3.33	7.09	9.46E-10	yes	72	SYNE2	2.31	8.25	9.16E-09	yes (DNS)
73	KIAA1212	2.34	10.05	3.07E-11	-	73	SYTL5	1.98	9.55	3.26E-09	yes
74	RAB31	2.77	8.48	3.01E-11	-	74	COL18A1	2.34	8.02	2.34E-10	-
75	FANCD2	3.41	6.87	1.02E-09	-	75	ISG20	2.15	8.65	2.53E-08	-
76	KIF11	3.32	7.04	1.33E-07	-	76	SYTL4	2.36	7.86	1.06E-05	-
77	BRCA2	3.90	5.99	1.08E-09	yes	77	SEPT9	1.72	10.71	1.81E-08	-
78	MCM2	2.64	8.84	7.18E-10	-	78	CCND1	1.83	9.95	2.55E-07	yes
79	HELLS	3.49	6.67	7.95E-11	-	79	SLC35E2	1.87	9.75	5.74E-09	-
80	MCM6	2.37	9.79	1.10E-09	yes	80	SLC25A24	1.72	10.58	2.16E-07	-
81	TTK	3.19	7.26	8.41E-09	-	81	MYO1B	1.69	10.73	6.67E-08	yes (DNS)
82	E2F7	3.94	5.88	6.39E-09	yes	82	FER1L3	1.63	11.04	9.86E-08	-
83	PRC1	2.42	9.55	4.49E-10	-	83	ADAMTS19	2.08	8.64	2.51E-10	-
84	RET	2.84	8.12	4.49E-10	-	84	ARHGAP26	2.84	6.29	8.64E-10	-
85	PSF1	2.77	8.31	2.48E-11	yes	85	KLF4	1.61	10.94	1.69E-07	-
86	XBP1	2.24	10.27	2.59E-09	yes	86	TPBG	1.55	11.38	4.87E-07	-
87	PLAC1	4.21	5.42	2.06E-11	-	87	HIPK2	2.13	8.25	2.16E-09	-
88	LOC441168	2.11	10.79	2.54E-10	-	88	NPY1R	2.01	8.71	2.56E-08	-
89	CDCA1	3.51	6.46	1.78E-07	-	89	EFEMP1	1.92	9.09	1.68E-08	-
90	HIST1H4C	2.29	9.88	6.50E-09	-	90	C6orf141	1.81	9.57	1.85E-08	-
91	RAD51	3.78	6.00	2.05E-11	-	91	TST	1.90	9.72	1.25E-08	-
92	KIF23	2.89	7.84	2.78E-07	-	92	MCM5	1.77	9.79	6.11E-07	-
93	PLK4	3.32	6.79	9.88E-09	yes (DNS)	93	GLA	1.90	9.06	4.33E-09	yes
94	FEN1	2.63	8.56	1.04E-10	-	94	THRAP4	2.51	6.84	6.29E-09	-
95	EGR3	3.95	5.70	1.92E-09	yes (DNS)	95	RBBP8	2.06	8.34	4.62E-07	yes (DNS)
96	DUT	2.36	9.52	4.49E-10	-	96	FLJ20366	2.61	6.57	6.62E-09	-
97	OFD1	2.02	11.10	5.46E-10	-	97	LAPTM4A	1.47	11.64	7.83E-07	-
98	POLE2	3.32	6.74	3.08E-10	-	98	SFRX2	2.01	8.51	1.39E-08	-
99	RNASEH2A	2.60	8.61	4.82E-10	yes	99	PCR	3.21	5.34	1.48E-07	yes
100	DHFR	2.86	7.78	3.23E-09	-	100	BRP4L	2.10	8.14	5.41E-09	-

Figure supplémentaire 1

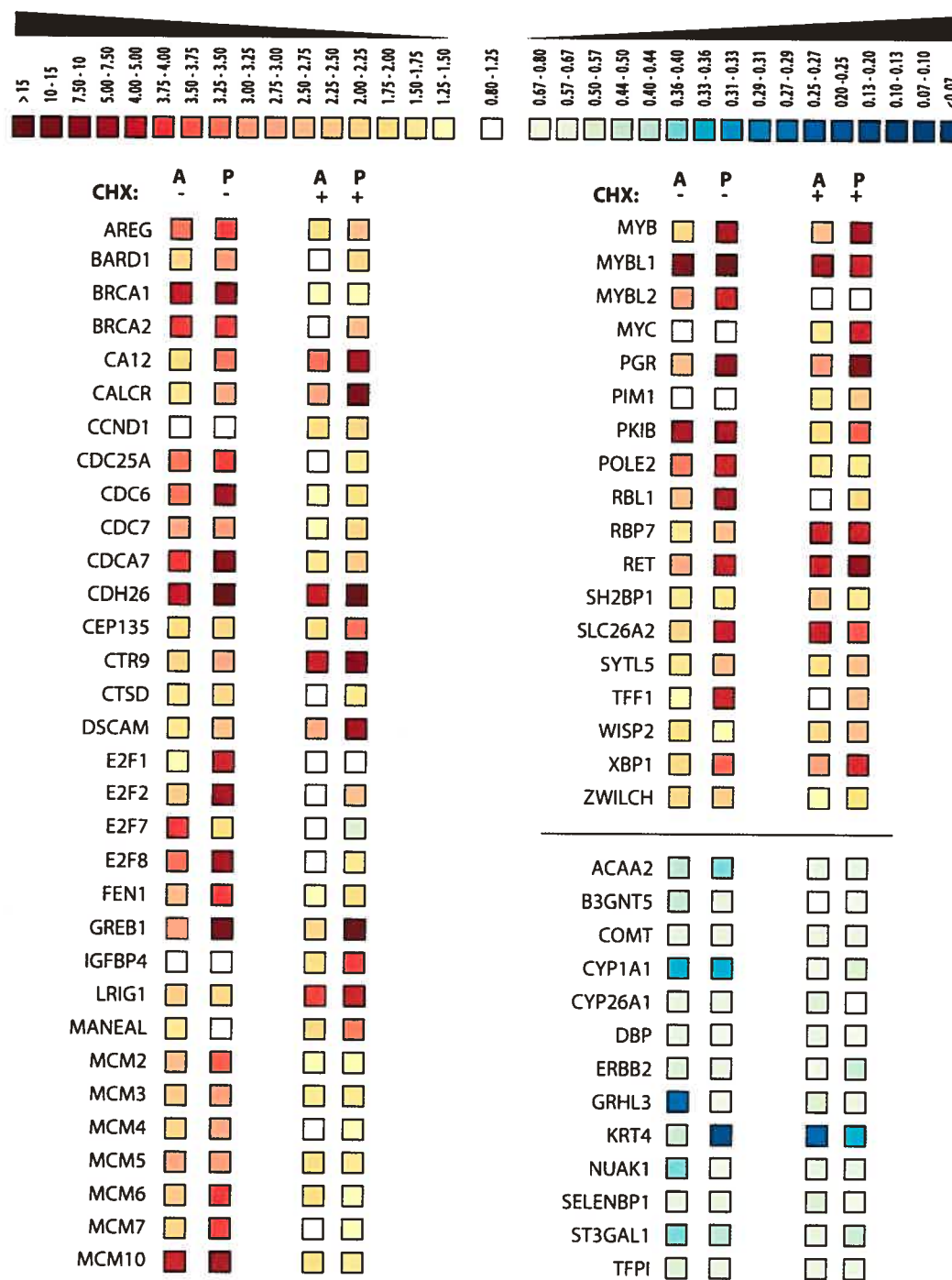


Figure supplémentaire 2

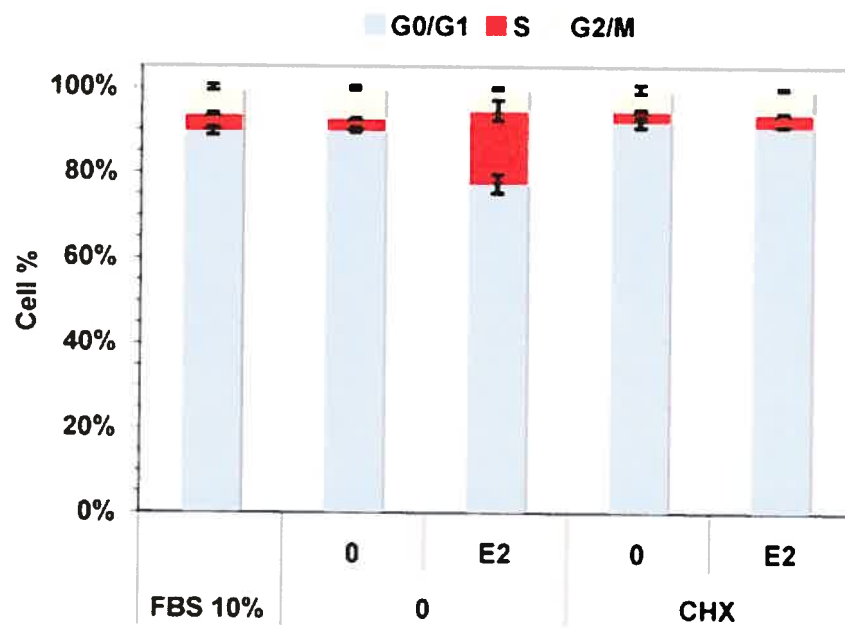


Table supplémentaire 1

UP													
Site	Matrix	TRANSFAC	Promoter	+/- 5 kb TSS					+/- 2.5 kb TSS				
				- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX	- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX
AHR/ARNT	ARNT_01	M00236	-TE										
CCAAT box	CAAT_01	M00254	-TE		1,70				1,67	2,04			1,54
	NFY_01	M00287	-TE	1,83	2,01				1,93	2,10			
CHOP-C/EBP	NFY_Q6	M00185	conserved							1,58			
	CHOP_01	M00249	-TE										
E2F	E2F_03	M00516		3,10	3,12	3,02			3,64	3,72	3,32		1,65
	E2F_03 core	M00516	conserved	4,44	4,09	5,70			5,48	5,05	7,04		1,64
EGR	EGR1_01	M00243	-TE										3,11
	EGR2_01	M00246	-TE					1,66			2,97		2,06
	EGR3_01	M00245	-TE					2,07			2,87		1,79
	NGFIC_01	M00244	-TE								3,63		2,26
ERE	ER_Q6	M00191	-TE				3,13	2,50			3,02		2,26
	ERE	-					2,53	2,57				3,84	2,93
ERE-derived	ER0	-	-TE				1,69	1,62				3,58	3,25
	ER5	-	-TE				1,84						
GC box	GC_01	M00255	-TE									2,34	
homeobox	CHX10_01	M00437	-TE			2,16						1,52	

DOWN													
Site	Matrix	TRANSFAC	Promoter	+/- 5 kb TSS					+/- 2.5 kb TSS				
				- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX	- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX
AP1	AP1_Q4	M00188	-TE							1,90			
	AP1_Q6	M00174	conserved						1,62	1,59			
C/EBP	CEBPB_02	M00117	conserved						2,67	3,12			
CCAAT box	NFY_Q6	M00185	conserved							1,53			
	FOX D3_01	M00130	conserved		1,69								
Forkhead	FOXJ2_01	M00422	conserved							1,53			
	FOXO4_01	M00472	conserved						1,54				
	FREAC3_01	M00291	-TE								2,11		2,07
	HFB8_01	M00294	conserved	1,52	1,50					1,52			
GC box	GC_01	M00255	-TE						1,68				
G-rich	MAZR_01	M00491	-TE							1,85			
	SPZ1_01	M00446	-TE									5,79	
homeobox	CHX10_01	M00437	-TE										2,23
NFKB	NFKAPPAB_01	M00054	-TE		1,71								
	NFKAPPAB65_01	M00052	-TE		1,66								
PAX	PAX4_02	M00377	-TE										1,70
	PAX4_03	M00378	conserved						2,00				
SP1	SP1_Q6	M00196	conserved		1,52				1,54	1,58			

Table supplémentaire 2

UP TOP 100													
Site	Matrix	TRANSFAC	Promoter	± 5 kb TSS					± 2.5 kb TSS				
				- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX	- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX
AP1	AP1_01	M00517					2,76						
cAMP-responsive element	CREB_01	M00039	-TE	2,13					2,47	2,47			
	CREB_Q4	M00178								1,89			
CCAAT box	CAAT_01	M00254	-TE						3,42	2,73			
	NFY_Q6	M00193										1,70	
	NFY_01	M00287	-TE	2,69	3,13				3,01	3,27			
	NFY_Q6	M00185	conserved		2,66				2,20				
CHOP-C/EBP	CHOP_01	M00249	-TE				3,71					5,01	
E2F	E2F_03	M00516		7,45	5,68				9,36	7,13			
	E2F_03 core	M00516	conserved	12,50	8,33				15,44	10,29			
EGR	EGR1_01	M00243	-TE			1,81					1,95		
	EGR3_01	M00245	-TE					2,30					
	NGFIC_01	M00244	-TE					2,88			1,99		
ERE	ER_Q6	M00191	-TE				8,67				2,39		
	ERE	-				4,00		5,00				4,91	5,89
ERE-derived	ERE_DR5	-					2,73						
	ERE_ER0	-					1,86						
	ERE_ER5	-	-TE				3,24					4,69	
	ERE_IR4	-					1,81						
Forkhead	HNF3B_01	M00131			1,55								
homeobox	CHX10_01	M00437	-TE			2,39							
Maf-recognition elements	BACH1_01	M00495					2,16						
PAR	HLF_01	M00260	-TE	3,36									
	VBP_01	M00228	-TE	3,70									

DOWN TOP 100													
Site	Matrix	TRANSFAC	Promoter	± 5 kb TSS					± 2.5 kb TSS				
				- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX	- CHX	Only - CHX	+ and - CHX	Only + CHX	+ CHX
C/EBP	CEBPB_01	M00109	conserved		1,86								
cAMP-responsive element	ATF_01	M00017	-TE		2,24					2,93			
	CREB_Q4	M00178								3,30			
CCAAT box	NFY_Q6	M00185	conserved		2,50					2,78			
E2F	E2F_03	M00516		2,30	3,55				2,46	3,57			
ERE-derived	ERE_DR6	-	-TE	1,80					2,33	1,91			
EVI	EVI1_06	M00011	conserved										2,21
Forkhead	FOXA	-							1,75	2,11			
	FOXO1_01	M00473	conserved		2,80								
	FOXO4_01	M00472	conserved	2,41	3,86				3,27	4,08			
GATA	GATA1_06	M00347	conserved		1,93								
G-rich	SPZ1_01	M00446										5,51	
homeobox	CHX10_01	M00437	-TE					3,17					3,82
Ikaros	IK1_01	M00086	-TE								2,81		
SOX	SRY_02	M00160	conserved					2,53					2,58
SP1	SP1_01	M00008	conserved		1,82				2,02	1,95			

Chapitre 7 : Regulation of GREB1 Transcription by Estrogen Receptor Alpha Through a Multipartite Enhancer Spread Over 20 KB of Upstream Flanking Sequences

Julie Deschênes¹, Véronique Bourdeau¹, John H. White, and Sylvie Mader²

Article accepté : *Journal of Biological Chemistry*, accelerated communication, JBC published April 26, 2007 as doi:10.1074/jbc.C700030200

À titre de co-premier auteure, j'ai mis au point les différentes techniques (ChIP, 3C, QPCR) dans le laboratoire et réalisé la majorité des expériences en collaboration étroite avec Dr Bourdeau. J'ai également participé à la rédaction du manuscrit et de sa correction

Institute for Research in Immunology and Cancer and Department of Biochemistry (J.D., V.B., S.M.), Université de Montréal, Montréal, Québec, Canada H3C 3J7; and Departments of Physiology (J.H.W.) and Medicine (J.H.W. and S.M.), McGill University, Montréal, Québec, Canada H3A 1A4. ¹ These authors contributed equally to this work. ² To whom correspondence should be addressed: IRIC, Univ. de Montréal, C.P.6128, Succursale Centre-Ville, Montréal, Qc, H3C 3J7, CANADA, Tel.:514-343-7166, Fax:514-343-7780; e-mail: sylvie.mader@umontreal.ca

ABSTRACT

Estrogen receptors activate transcription in part through direct interactions with specific DNA motifs, called estrogen response elements (EREs). Here we show that the strong and sustained induction of the gene regulated in breast cancer 1 (GREB1), a gene of unknown function that has been previously suggested to play a role in the effects of estradiol on breast cancer cell proliferation (Rae JM et al., 2005, *Breast Cancer Res Treat* 92:141-9), is mediated by binding of estrogen receptor alpha (ER α) to three consensus EREs spread over ~20 kb of upstream flanking sequences. In addition to ER α , coactivator SRC-3, acetylated histones and phosphorylated RNA polymerase II (P-polII) were detected on all three EREs in the presence of estrogen, while basal recruitment of ER α and P-polII was observed only on the proximal element. Chromatin loops were formed between each ERE and the GREB1 transcriptional start site in the presence of estrogen, but not of a total antiestrogen. Furthermore, estradiol induced physical association between EREs, suggesting that these elements function as a potent multi-partite enhancer to regulate GREB1 transcription.

INTRODUCTION

Estrogens are steroid hormones that have pleiotropic physiological actions in numerous target tissues including the reproductive system (1-3), but also play an important role in cancers, in particular breast cancer (4,5). Estrogens act through two receptors, ER α and ER β , members of the nuclear receptor superfamily of ligand-inducible transcription factors (6,7). ER α is expressed or overexpressed in two thirds of breast tumors (8), and blockade of estrogen signaling through antiestrogens, which are competitive inhibitors of ERs, is an effective treatment of the majority of ER α -positive breast tumors (9,10). Functional genomic approaches have led to the identification of a variety of estrogen target genes in ER α -positive breast cancer cell lines (11-17). Some estrogen target genes, such as those encoding cyclin D1 (CCND1), c-myc (MYC), and E2F transcription factors have been shown to mediate the proliferative effects of estrogen in breast cancer cell lines (18-21). It has also been recently suggested that the gene regulated in breast cancer 1, GREB1 (22), whose function remains unknown, contributes to the growth-promoting effects of estrogens in MCF7 cells (23).

ERs can bind to specific DNA motifs, called estrogen response elements (EREs) through a central conserved DNA binding domain (24). Consensus EREs, which were defined by compiling natural response elements in estrogen-responsive genes and correspond to the highest affinity binding site *in vitro*, are 15 bp-palindromes composed of two PuGGTCA motifs spaced by 3 bps (25,26). Although most high-affinity EREs were initially characterized in the proximal regulatory sequences of target genes, recent studies have indicated that elements binding ER α can be found at large distances from transcriptional start sites (TSS). Our genome-wide mapping of high affinity EREs combined with characterization of elements by chromatin immunoprecipitation identified functional ER α binding sites up to 10 kb from TSS (27). In addition, genome-wide mapping of chromatin fragments bound by ER α revealed sites located much further from adjacent genes, and use of chromatin conformation capture assays showed that some of

these elements can form chromatin loops with promoters located at distances of up to 100 kb (15,17). Often, ER α binds several chromatin regions within this range of distances from the TSS of estrogen-regulated genes, raising the possibility of cooperativity between widely separated enhancer units for transcriptional regulation of target genes. This may result in the formation of multiple chromatin loops with the TSS of these target genes, and possibly between enhancers themselves.

GREB1 stands out as an estrogen target gene because of the presence in its flanking region of three consensus EREs spread over 20 kb of upstream sequences. We have previously observed *in vivo* recruitment of ER α to the EREs present at -1.5 and -9.5 kb (27). Here we address whether the three GREB1 consensus EREs are functional enhancers that cooperate for transcriptional induction of the GREB1 gene by ER α , and present evidence for a complex chromatin loop structure that implicates all three EREs and the TSS.

EXPERIMENTAL PROCEDURES

Cell culture MCF7 breast carcinoma cells and Ishikawa cells were maintained in α -minimal Eagle's medium (α -MEM) (Wisent, St-Bruno, QC, Canada) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, Oakville, On, Canada). ZR75 and T47D were maintained in RPMI (Wisent) supplemented with 10% FBS. MDA-MB-231::ER α cells were maintained in DMEM (Wisent) supplemented with 5% FBS with 0.15 mg/ml of hygromycin (Sigma). Three days before experiments, cells were switched to phenol red-free Dulbecco's modified Eagle's medium (DMEM) (Wisent) containing 10% charcoal-stripped FBS (FBS-T), 1% sodium pyruvate (Wisent), 1% penicillin/streptomycin (Wisent) and 1% L-glutamine (Wisent). The day before hormonal stimulation, the medium was changed to phenol red-free DMEM supplemented with 0.5% charcoal-treated FBS. Cells were then treated with 17- β -estradiol (E2, 25 nM, Sigma), ICI 182,780 (ICI, 100 nM,

Tocris, Ellisville, Missouri, USA) or vehicle (ethanol) for variable periods of time as indicated in figure legends.

RNA extraction Cells were seeded in 10 cm plates (MCF7, T47D, ZR75, MDA-MB-231::ER α) or 6 cm plates (Ishikawa) at a density such that near confluency was obtained at the end of the treatment. Cells were treated with 25 nM E2, 100 nM ICI 182,780 or vehicle (ethanol) for different times, as indicated in figure legends. For pre-treatments with actinomycin D (2 μ g/ml) or cycloheximide (10 μ g/ml), incubation was initiated 1 h before hormonal treatment. siGenome siRNAs were transfected according to the instructions of the manufacturer (Dharmacon, Chicago, IL, USA), medium was changed 24 h after transfection and hormonal stimulations were initiated after another 24 h. At the end of the treatments, medium was removed, cells were collected in 1 ml of TRI Reagent (Sigma) and total RNA was extracted as recommended by the manufacturer.

Reverse transcription and real-time PCR Total RNA (2 μ g) was reverse transcribed using the RevertAid H first minus strand cDNA synthesis kit (MBI Fermentas, Burlington, On, Canada) as recommended by the manufacturer. The RT product was diluted 10 times prior to real-time PCR. Each real-time PCR amplification reaction contained the RT dilution (2 μ l), forward and reverse primers (150 to 300 nM), MgCl₂ (3 to 4.5 nM according to primer pairs), dNTP (0.2 mM), SYBR Green (0.33X, Invitrogen, Burlington, On, Canada), buffer for Jump Start Taq and Jump Start Taq (0.5 U, Sigma) in a final volume of 20 μ l. After denaturation at 95°C for 7 min, samples went through a one-degree annealing temperature touch down of 7 cycles starting from 60°C (30 s at 95°C, 30 s at annealing temperature, 30 s at 72°C) followed by 40 cycles of amplification (30 s at 95°C, 30 s at 58°C and 30 s at 72°C). A dissociation protocol followed the amplification program to characterize the amplified product(s). PCR was performed using a RotorGene 3000 (Corbett, Australia) and analyzed using expression levels of the p36b4 gene for normalization. For each set of primers, non-template control reactions were performed as a negative control. Each sample was assayed in triplicate and each experiment was

reproduced at least two times. A typical experiment is shown. All primer pairs were designed using the Primer3 software and chosen to span an intron. The primer sequences and individual conditions used for polymerase chain reaction amplification are available upon request.

Western blot analysis Whole cell extracts were prepared as described previously (28) using anti-ER α mouse monoclonal (B10, kind gift from Prof. P. Chambon) and anti- β -actin mouse monoclonal antibody (Sigma Diagnostics).

Chromatin immunoprecipitation For chromatin immunoprecipitation (ChIP) assays, cells in phenol red-free DMEM containing 0.5% FBS-T were treated with vehicle, E2 or ICI for 45 min. Chromatin was cross-linked by treating cells with 1.5% formaldehyde for 10 min at room temperature and fragmented by sonication as previously reported (27), yielding fragments of average size approximately 400 bp. Antibodies against hER α (SC-543), SRC-3 (SC-7217) or β -actin were purchased from Santa Cruz Biotechnology, Inc (Santa Cruz, CA, USA) and antibodies against phosphorylated polymerase II (05-623) or acetylated H4 (06-598) were purchased from Upstate Biotechnology, Inc (Chicago, IL, USA). The sequences of the primers used in ChIP assays (Sigma Genosis) are available upon request. Chromatin immunoprecipitation experiments were performed three times with similar results. A representative set of results is shown.

Chromatin Conformation Capture (3C) assays 3 C assays were performed essentially as described (29), with only minor modifications. In short, three days prior to the experiment, MCF-7 cells were switched to phenol red-free DMEM containing 10% FBS-T for 48 h (9 million cells per 15 cm petri dish). The serum concentration was decreased to 0.5% 24 h before treatment with E2 (25 nM), ICI (100 nM) or vehicle (ethanol) for 45 min. Culture medium was removed and cells were fixed with 1.5% formaldehyde for 10 min at room temperature. Cell were then washed twice with cold phosphate buffer saline (PBS) solution, collected in 500 μ l collection buffer (100 mM Tris-

HCl pH 9.4, 10 mM DTT and protease inhibitor cocktail), washed with cold PBS and resuspended in ice-cold lysis buffer (10 mM Tris-HCl pH8.0, 10 mM NaCl, 0.2% NP-40 and protease inhibitor cocktail). Nuclei were resuspended in 2 ml Tango 1.2X buffer (MBI Fermentas) supplemented with SDS. Triton X-100 was added to sequester the SDS. The crosslinked DNA was digested overnight with 400 U of restriction enzyme MspI or XceI (MBI Fermentas), and then with another 100 U for 2 h at 37°C. The restriction enzyme was inactivated by addition of SDS and incubation at 65°C. Genomic DNA concentrations were measured in 30 µl aliquots for standardization between samples. The reactions were diluted with ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP and 25 µg/ml BSA), supplemented with Triton X-100 (1% final concentration) and incubated for 1 h at 37°C. The DNA was ligated using T4 ligase (NEB, Ipswich, MA, USA) overnight at 16°C. RNase was added, and samples were incubated overnight at 65°C to reverse the crosslinks. The following day, samples were incubated for 2 h at 45°C with proteinase K, and the DNA was purified by phenol-chloroform extractions and ethanol precipitation. Chromatin loop formation was assessed by PCR amplification carried out using similar conditions as for Q-PCR amplification but without SYBR Green, and with two nested primer pairs for each predicted ligation event. PCR products were visualized on ethidium bromide agarose gels (2%) and confirmed by sequencing. Primer sequences are available upon request.

RESULTS AND DISCUSSION

To better characterize the mechanisms of regulation of the GREB1 gene by estrogen, we first assessed the influence of estradiol (E2) treatment on its expression in different ER⁺ cell lines (MCF7, T47D, ZR45, MDA-MB-231::ER α and Ishikawa). Transcriptional induction was detected in all ER⁺ cells with E2 treatment whereas the antiestrogen ICI 182,780 (ICI) reduced mRNA levels (Suppl. Fig. 1A). Consistent with these findings, a strong correlation between higher GREB1 mRNA levels and ER α

expression in breast tumors can be observed in various microarray studies compiled by Oncomine (www.oncomine.org; Suppl. Fig. 1B).

To confirm the involvement of ER α in transcriptional regulation by estrogen, we assessed the effect of siRNA-mediated suppression of ER α on GREB1 mRNA expression levels in MCF7 cells. As shown in Fig. 1, levels of ER α proteins were reduced by the specific siRNA duplex (Fig. 1A) resulting in a substantial reduction of both basal and estrogen-stimulated levels of GREB1 mRNA (Fig. 1B), which together with the effect of ICI treatment indicates that ER α is important for both basal and activated GREB1 mRNA expression. To further verify that GREB1 is regulated at the transcriptional level by E2, MCF7 cells were pre-treated by the transcription inhibitor actinomycin D. This treatment prevented induction of GREB1 transcripts, indicating that E2 enhances GREB1 *de novo* transcription (Fig. 1C). On the other hand, the protein synthesis inhibitor cycloheximide (CHX) had no effect on the induction by E2 in MCF7 cells as reported (22), similar to what was observed with the E2 target gene encoding the progesterone receptor (PGR, Suppl. Fig. 2A), while induction of the BRCA1 gene was abrogated by CHX and that of pS2/TFF1 was largely attenuated as previously reported (30) (Suppl. Fig. 2A). Kinetics of induction of GREB1 transcripts by E2 in MCF7 cells resembled that of the PGR target gene, being both rapid (detectable at 2 h), and sustained over a 48 h period (Suppl. Fig. 2B). On the other hand, induction of E2 targets MYC or FOS was more rapid but transient (Suppl. Fig. 2C), while that of pS2/TFF1 was progressive and that of BRCA1 delayed in time (Suppl. Fig. 2D). Together, these results indicate that GREB1 is a primary ER α target induced in an early, strong and sustained manner by ER α in MCF7 cells.

The 5' flanking sequences of GREB1 contain three consensus EREs (Fig. 2A), and we have previously shown that two of the three GREB1 EREs strongly bound ER α in ChIP experiments in MCF7 cells and in MDA-MB-231 cells stably transfected with ER α (27). Recruitment of ER α to the three EREs was also observed in a recent genome-wide mapping of ER α binding sites in the presence of estrogen (17). Here we show that in

MCF7 cells all three perfect EREs are bound with a cyclical temporal pattern after E2 stimulation (Fig. 2B), as previously observed for the pS2/TFF1 gene (31,32). Note, however, that the basal levels of ER α recruited to the ERE at -1.5 kb from the GREB1 TSS were higher than those for other EREs. Only faint binding of ER α to the GREB1 TSS was observed under our experimental conditions. In addition, recruitment to the EREs of the ER co-activator SRC-3 (Fig. 2C), acetylated histone H4 (Fig. 2D) as well as phosphorylated RNA polymerase II (P-polII) (Fig. 2C) was also detected by ChIP assays. Note that binding of P-polII was strong under basal conditions on the -1.5 kb ERE. Altogether, these data suggest that the two distal EREs are functional enhancers and that the proximal ERE contributes to the low basal, ICI-repressed activity of the promoter.

While ChIP data indicate recruitment of ER α and its coactivator SRC-3, acetylation of H4 and binding of P-polII to sites that are closer to the GREB1 TSS than to any other gene (the closest gene, E2F6, is located more than 45 kb away from ERE3), this does not conclusively demonstrate that these EREs play a role in direct transactivation of GREB1. To investigate whether all three EREs associate with the GREB1 TSS, we used a chromatin conformation capture (3C) assay (33), which detects physical proximity between distal DNA sites by ligation of restricted DNA fragments that are maintained juxtaposed by crosslinked DNA/protein complexes. Using restriction enzymes that cut at least once between consensus EREs and the GREB1 TSS, we observed PCR amplification of products resulting from the ligation of each of the three EREs and the TSS (Fig 3A). Association of ERE1 (-1.5 kb) and the TSS was constitutive, which could reflect either the higher basal level of receptor and P-polII associated with this ERE (see Fig. 2C) or the short distance between the ERE and the TSS. However, the lack of chromatin loop formation in the presence of ICI suggests formation of a chromatin loop in the absence of hormone rather than random interactions between the two DNA fragments. No chromatin loops were detected with several regions located between the EREs and the TSS (Fig 3A and data not shown) or between the TSS and a fragment at +1.5 kb downstream (data not shown), demonstrating that E2-induced chromatin looping with the TSS is specific for fragments

containing the EREs. Note that the 3C PCR products obtained were sequenced to confirm product identity and that all primer sets for 3C detection led to amplification of a digested and re-ligated BAC clone overlapping this genomic region (data not shown). These results indicate that the consensus EREs in the GREB1 gene behave as E2-induced enhancers since they recruit ER α , SRC-3 and P-polII, are associated with acetylated histones and interact with the TSS of GREB1 in the presence of E2.

The presence of three strong enhancers raises the question of whether these sequences associate with the TSS in a mutually exclusive and possibly sequential manner, or if they may participate in reciprocal interactions as well. To investigate this question, we analyzed by 3C whether chromatin loop formation could take place between ERE enhancers themselves (Fig 3B). PCR amplification resulting from chromatin loop formation between any combinations of two consensus EREs could be detected in the presence of E2 but not in its absence or in the presence of ICI. In addition, no interactions were detected with other intermediate regions (Fig. 3B and data not shown), demonstrating again the specificity of these interactions for the presence of EREs. These results suggest that EREs can physically associate with each other, either in a large complex with the TSS or in transient binary interactions (see model in Suppl. Fig. 3).

Based on the above findings, we conclude that the 5'-flanking region of GREB1, a gene of as yet unknown function but which may play an important role in E2-induced cell proliferation (23), contains a strong, multipartite enhancer composed of three consensus EREs separated by 8 to ~20 kb of intervening sequences. Future studies will be necessary to assess whether this multipartite enhancer also interacts with additional ER α binding regions and with the TSS of other E2 target genes, and whether transcriptional regulation by E2 often involves the association of several widely-spread enhancer units

REFERENCES

1. Couse, J. F., and Korach, K. S. (1999) *Endocrine Rev* **20**, 358-417
2. McEwen, B. S., and Alves, S. E. (1999) *Endocrine Rev* **20**, 279-307
3. Jordan, V. C. (2001) *J Natl Canc Inst* **93**, 2-4
4. Pike, M. C., Spicer, D. V., Dahmouch, L., and Press, M. F. (1993) *Epidemiol Rev* **15**(1), 17-35
5. Simpson, E. R. (2003) *J Steroid Biochem Mol Biol* **86**(3-5), 225-230
6. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995) *Cell* **83**(6), 835-839
7. Robinson-Rechavi, M., Escriva Garcia, H., and Laudet, V. (2003) *J Cell Sci* **116**(Pt 4), 585-586.
8. Ali, S., and Coombes, R. C. (2002) *Nat Rev Cancer* **2**(2), 101-112
9. O'Regan, R. M., and Jordan, V. C. (2002) *Lancet Oncol* **3**(4), 207-214
10. Howell, A., Robertson, J. F., Abram, P., Lichinitser, M. R., Elledge, R., Bajetta, E., Watanabe, T., Morris, C., Webster, A., Dimery, I., and Osborne, C. K. (2004) *J Clin Oncol* **22**(9), 1605-1613
11. Soulez, M., and Parker, M. G. (2001) *J. Mol. Endocrinol.* **27**, 259-274
12. Frasor, J., Danes, J. M., Komm, B., Chang, K. C., Lyttle, C. R., and Katzenellenbogen, B. S. (2003) *Endocrinology* **144**(10), 4562-4574
13. Lin, C. Y., Strom, A., Vega, V. B., Kong, S. L., Yeo, A. L., Thomsen, J. S., Chan, W. C., Doray, B., Bangarusamy, D. K., Ramasamy, A., Vergara, L. A., Tang, S., Chong, A., Bajic, V. B., Miller, L. D., Gustafsson, J. A., and Liu, E. T. (2004) *Genome Biol* **5**(9), R66
14. Cicatiello, L., Scafoglio, C., Altucci, L., Cancemi, M., Natoli, G., Facchiano, A., Iazzetti, G., Calogero, R., Biglia, N., De Bortoli, M., Sfiligoi, C., Sismondi, P., Bresciani, F., and Weisz, A. (2004) *J Mol Endocrinol* **32**(3), 719-775

15. Carroll, J. S., Liu, X. S., Brodsky, A. S., Li, W., Meyer, C. A., Szary, A. J., Eeckhoutte, J., Shao, W., Hestermann, E. V., Geistlinger, T. R., Fox, E. A., Silver, P. A., and Brown, M. (2005) *Cell* **122**(1), 33-43
16. Vega, V. B., Lin, C. Y., Lai, K. S., Kong, S. L., Xie, M., Su, X., Teh, H. F., Thomsen, J. S., Yeo, A. L., Sung, W. K., Bourque, G., and Liu, E. T. (2006) *Genome Biol* **7**(9), R82
17. Carroll, J. S., Meyer, C. A., Song, J., Li, W., Geistlinger, T. R., Eeckhoutte, J., Brodsky, A. S., Keeton, E. K., Fertuck, K. C., Hall, G. F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E. A., Silver, P. A., Gingeras, T. R., Liu, X. S., and Brown, M. (2006) *Nat Genet* **38**(11), 1289-1297
18. D'Cruz, C. M., Gunther, E. J., Boxer, R. B., Hartman, J. L., Sintasath, L., Moody, S. E., Cox, J. D., Ha, S. I., Belka, G. K., Golant, A., Cardiff, R. D., and Chodosh, L. A. (2001) *Nat Med* **7**(2), 235-239
19. Hui, R., Finney, G. L., Carroll, J. S., Lee, C. S., Musgrove, E. A., and Sutherland, R. L. (2002) *Cancer Res* **62**(23), 6916-6923
20. Butt, A. J., McNeil, C. M., Musgrove, E. A., and Sutherland, R. L. (2005) *Endocr Relat Cancer* **12 Suppl 1**, S47-59
21. Mukherjee, S., and Conrad, S. E. (2005) *J Biol Chem* **280**(18), 17617-17625
22. Ghosh, M. G., Thompson, D. A., and Weigel, R. J. (2000) *Cancer Res* **60**(22), 6367-6375.
23. Rae, J. M., Johnson, M. D., Scheys, J. O., Cordero, K. E., Larios, J. M., and Lippman, M. E. (2005) *Breast Cancer Res Treat* **92**(2), 141-149
24. Mader, S., Chambon, P., and White, J. H. (1993) *Nucleic Acids Res* **21**(5), 1125-1132
25. Klinge, C. M. (2001) *Nucleic Acids Res* **29**(14), 2905-2919
26. Sanchez, R., Nguyen, D., Rocha, W., White, J. H., and Mader, S. (2002) *Bioessays* **24**(3), 244-254

27. Bourdeau, V., Deschenes, J., Metivier, R., Nagai, Y., Nguyen, D., Bretschneider, N., Gannon, F., White, J. H., and Mader, S. (2004) *Mol Endocrinol* **18**(6), 1411-1427
28. Rocha, W., Sanchez, R., Deschenes, J., Auger, A., Hebert, E., White, J. H., and Mader, S. (2005) *Mol Pharmacol* **68**(6), 1852-1862
29. Splinter, E., Grosveld, F., and de Laat, W. (2004) *Methods Enzymol* **375**, 493-507
30. Marks, J. R., Huper, G., Vaughn, J. P., Davis, P. L., Norris, J., McDonnell, D. P., Wiseman, R. W., Futreal, P. A., and Iglehart, J. D. (1997) *Oncogene* **14**(1), 115-121
31. Shang, Y., Hu, X., DiRenzo, J., Lazar, M. A., and Brown, M. (2000) *Cell* **103**, 843-852
32. Metivier, R., Penot, G., Hubner, M. R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003) *Cell* **115**(6), 751-763
33. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002) *Science* **295**(5558), 1306-1311

FOOTNOTES

The authors thank Drs Raphael Metivier and Michael Witcher for discussions on the ChIP technique and Drs Wouter de Laat, Hugo Wurtele, Pierre Chartrand and Josée Dostie for advice on the 3C technique. We are also grateful to Marie-France Proteau for technical assistance and Walter Rocha for helpful discussions. JD was recipient of studentships from the Fonds de la Recherche en Santé du Québec and the Faculté des Études Supérieures de l'Université de Montréal, and VB was holder of a post-doctoral fellowship from the Canadian Institutes for Health Research (CIHR). JHW and SM are Chercheurs-Boursiers Nationaux of the Fonds de Recherche en Santé du Québec (FRSQ). SM is the holder of the CIBC Breast Cancer Research Chair at Université de Montréal. This work was supported by CIHR grant MOP-13147 to SM.

FIGURE LEGENDS

Figure 1 : GREB1 is transcriptionally induced by ER α .

A. siRNA-mediated depletion of ER α protein levels was verified by western analysis with antibodies against ER α or β -actin 72 h after transfection. **B.** siRNA-mediated depletion of ER α in MCF7 cells reduces basal and E2-stimulated mRNA levels of GREB1. MCF7 cells were transfected by siRNA against ER α or GAPDH for 72 h and stimulated with hormone during the last 24 h prior to RNA extraction. **C.** The transcriptional inhibitor actinomycin D abrogates the induction of GREB1 mRNA levels by E2. Cells were pre-treated with actinomycin D (2 μ g/ml) for one hour before treatment with E2 for 4 h.

Figure 2: Consensus EREs recruit ER α , SRC-3, acetylated histones and phosphorylated RNA polymerase II in the presence of estradiol.

A. Position of the three consensus EREs (cEREs) in the GREB1 upstream regulatory regions. **B.** Cyclical recruitment of ER α on the three consensus EREs in GREB1 promoter. Chromatin immunoprecipitation (ChIP) was performed using an antibody against ER α or IgG as negative control with MCF7 cells treated with E2 (25 nM) or vehicle (ethanol) for

different times (as indicated). **C.** ChIP of SRC-3 and P-PolII at 2 h. **D.** ChIP of acetylated H4 at 2 h. Results are representative of at least two experiments.

Figure 3: Induction by estradiol treatment of chromatin loops associating the GREB1 enhancers and transcriptional start site.

A. Detection of chromatin loops formed between different regions of the GREB1 upstream sequences and the most 5' transcriptional start site. Chromatin Conformation Capture (3C) assays were performed in MCF7 cells treated with E2 (25 nM), ICI (100 nM) or vehicle (ethanol) for 45 min. **B.** Detection of chromatin loops formed between consensus EREs. Experiments are representative of three to four independent experiments performed either with the MspI or the XceI enzyme. Each chromatin loop was observed in between 2-4 independent assays.

Supplementary Figure 1: GREB1 mRNA is induced by estradiol in ER α -positive cancer cells

A. Induction of GREB1 mRNA by estradiol in different ER+ cancer cell lines. The different cell lines were treated with 17 β -estradiol (E2, 25 nM), ICI 182,720 (ICI, 100 nM) or vehicle (0, ethanol) and harvested at 24 h. Expression levels of GREB1 mRNA were assessed by Q-PCR amplification after RNA extraction and reverse transcription. Experiments were performed at least twice with similar results and a representative experiment is shown. **B.** GREB1 and ER α expression are correlated in ER-/ER+ breast tumor cells. Box plot graphs are taken from www.oncomine.org, 1) P value: 1.1E-34 (34), 2) P-value: 7.3E-20 (35), 3) P-value: 7E-17 (36), 4) P-value: 9.1E-12 (37), 5) P-value: 1.8E-8 (38), 6) P-value: 2.2E-6 (39).

Supplementary Figure 2: The strong and sustained induction of GREB1 by estradiol is a primary transcriptional event.

A. Induction of GREB1 is observed in the absence or presence of protein synthesis inhibitors. Cells were pre-treated with cycloheximide (cyclo, 10 μ g/ml) for one hour before treatment with E2 or ICI for 24 h. Other estrogen target genes were also analyzed in parallel using the same RNA samples. **B-D.** Induction of

GREB1 by E2 is both rapid and sustained. Fold induction over expression in the presence of vehicle were measured at different times for GREB1 and other estrogen target genes. Experiments were performed three times with similar results, and a representative experiment is shown

Supplementary Figure 3: The consensus GREB1 EREs function as a multipartite enhancer.

Model illustrating the possible effect of estrogen on chromatin loop formation between EREs and the transcriptional start site of GREB1. Note that these interactions may be simultaneous or occur in a sequential manner.

SUPPLEMENTARY REFERENCES:

34. van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E. T., Friend, S. H., and Bernards, R. (2002) *N Engl J Med* **347**(25), 1999-2009
35. Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M. E., Yu, J., Jatkoe, T., Berns, E. M., Atkins, D., and Foekens, J. A. (2005) *Lancet* **365**(9460), 671-679
36. Hess, K. R., Anderson, K., Symmans, W. F., Valero, V., Ibrahim, N., Mejia, J. A., Booser, D., Theriault, R. L., Buzdar, A. U., Dempsey, P. J., Rouzier, R., Sneige, N., Ross, J. S., Vidaurre, T., Gomez, H. L., Hortobagyi, G. N., and Pusztai, L. (2006) *J Clin Oncol* **24**(26), 4236-4244
37. Richardson, A. L., Wang, Z. C., De Nicolo, A., Lu, X., Brown, M., Miron, A., Liao, X., Iglehart, J. D., Livingston, D. M., and Ganesan, S. (2006) *Cancer Cell* **9**(2), 121-132

38. Zhao, H., Langerod, A., Ji, Y., Nowels, K. W., Nesland, J. M., Tibshirani, R., Bukholm, I. K., Karesen, R., Botstein, D., Borresen-Dale, A. L., and Jeffrey, S. S. (2004) *Mol Biol Cell* **15**(6), 2523-2536
39. Miller, L. D., Smeds, J., George, J., Vega, V. B., Vergara, L., Ploner, A., Pawitan, Y., Hall, P., Klaar, S., Liu, E. T., and Bergh, J. (2005) *Proc Natl Acad Sci U S A* **102**(38), 13550-13555

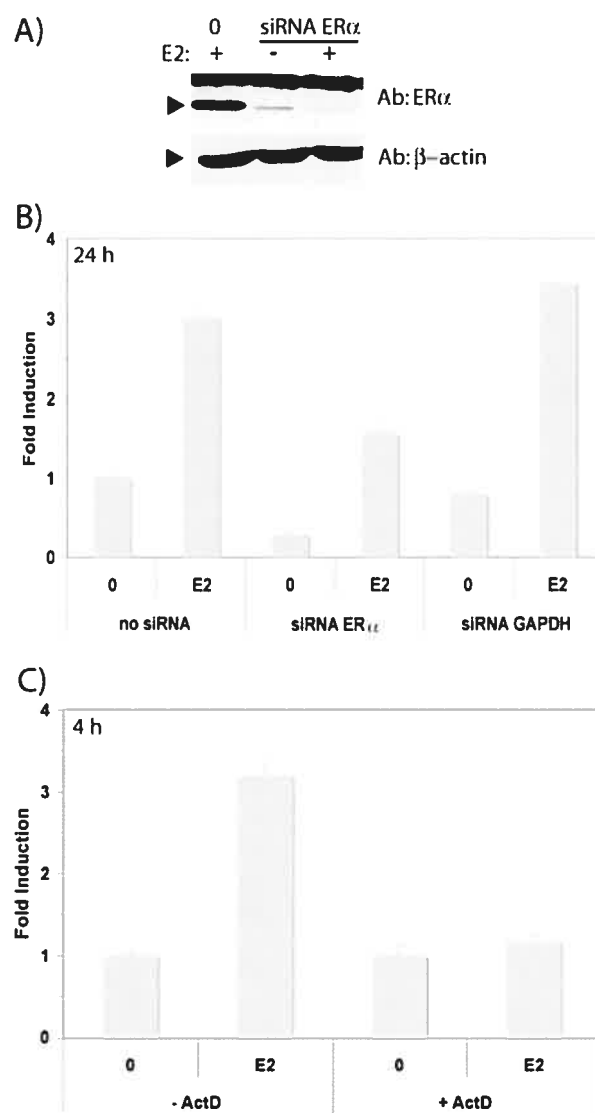
Figure 1

Figure 2

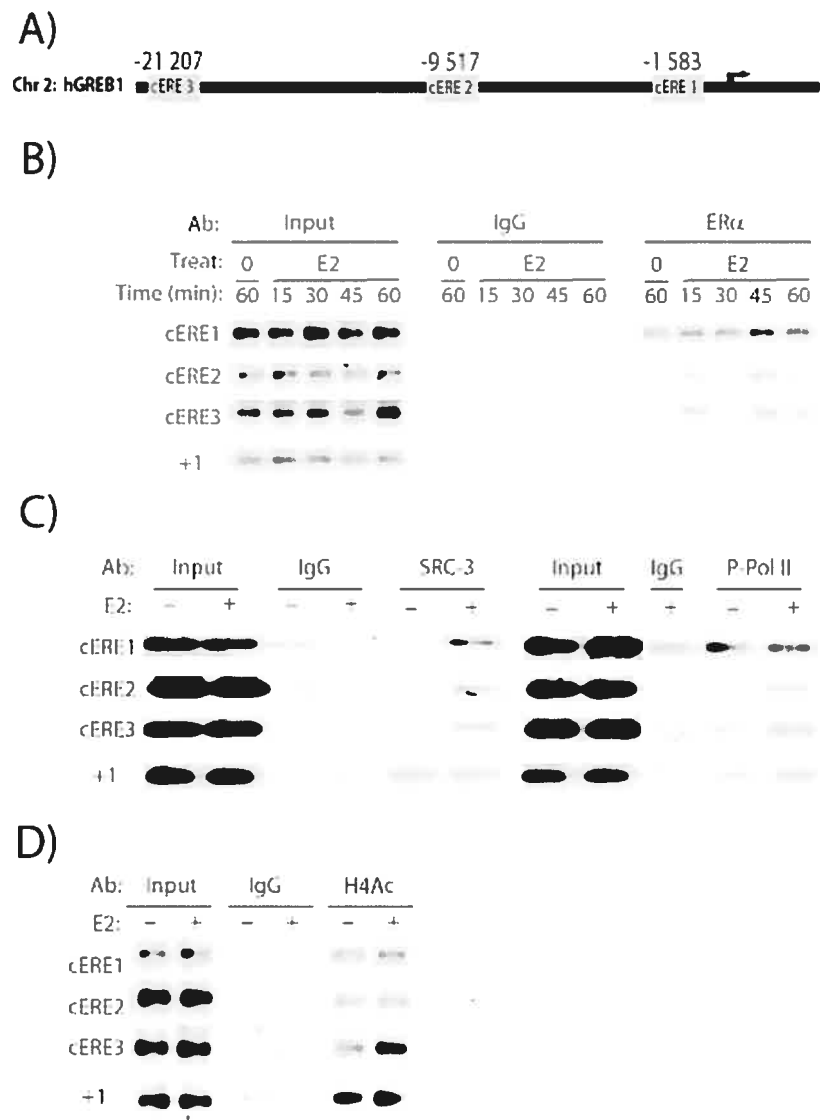


Figure 3

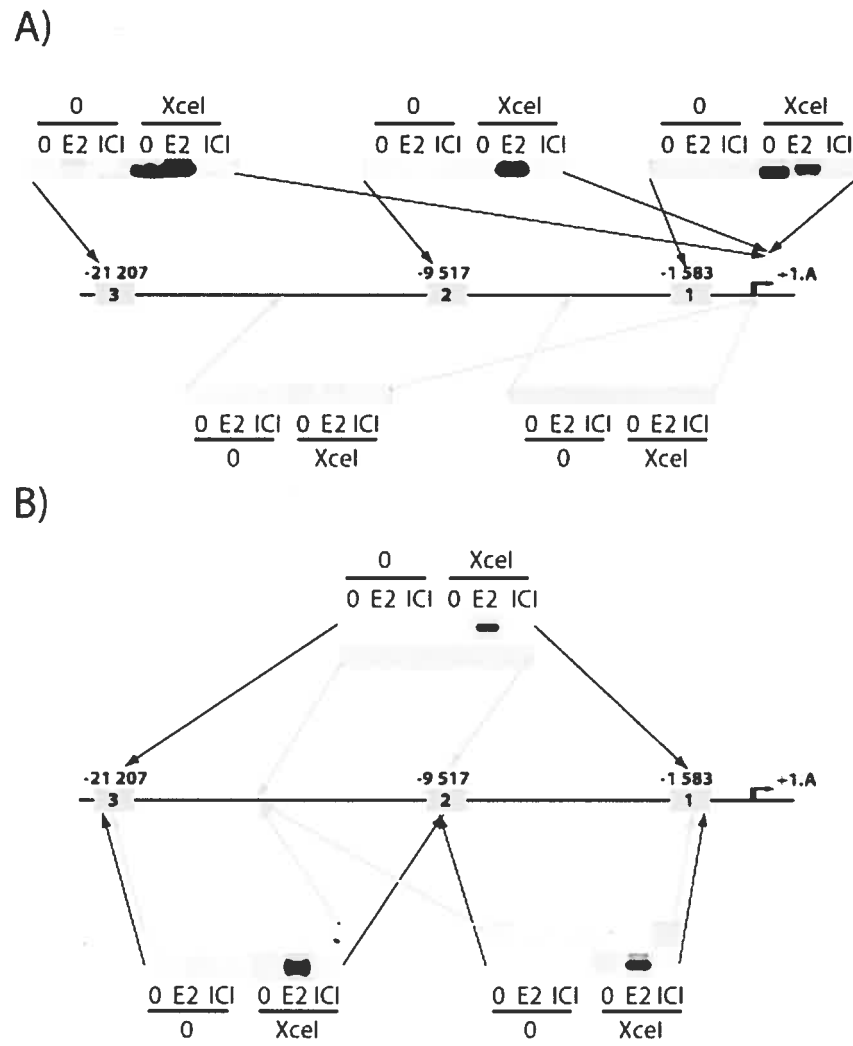


Figure Supplémentaire 1

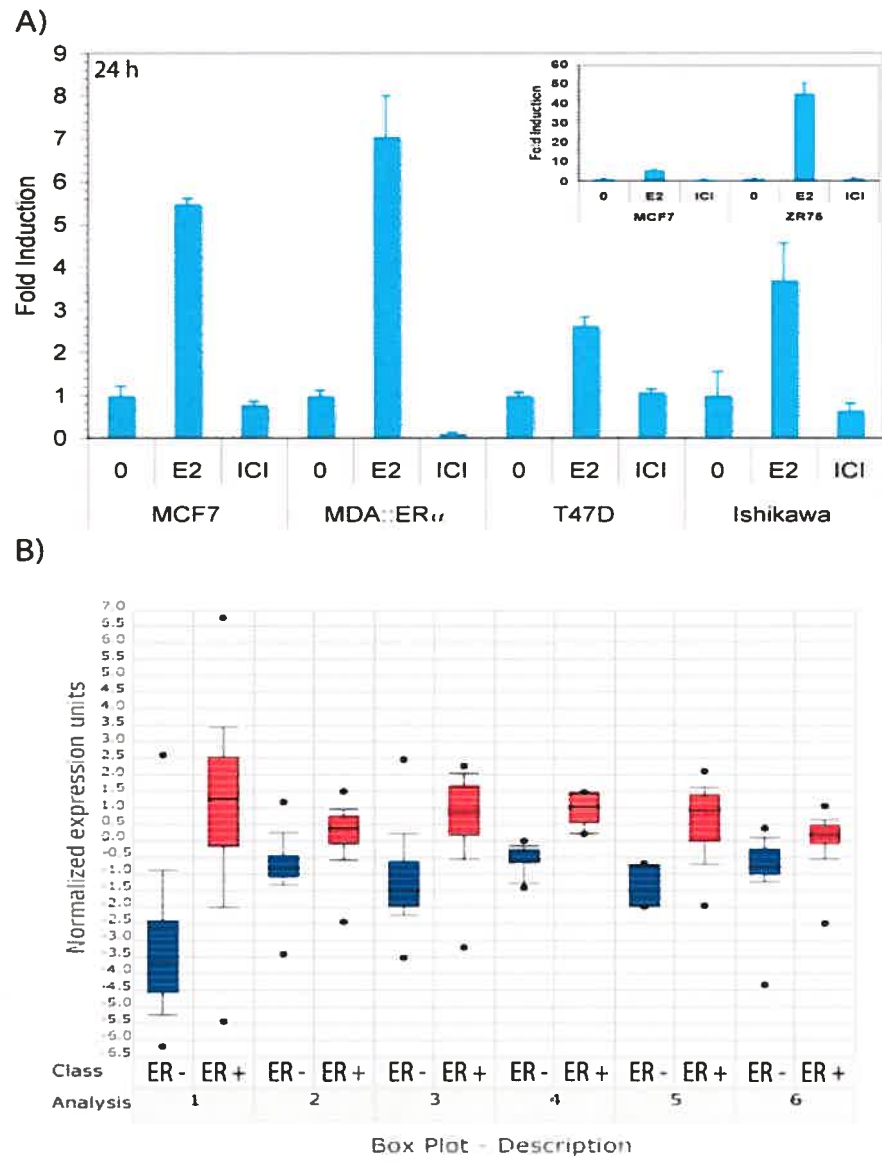


Figure Supplémentaire 2

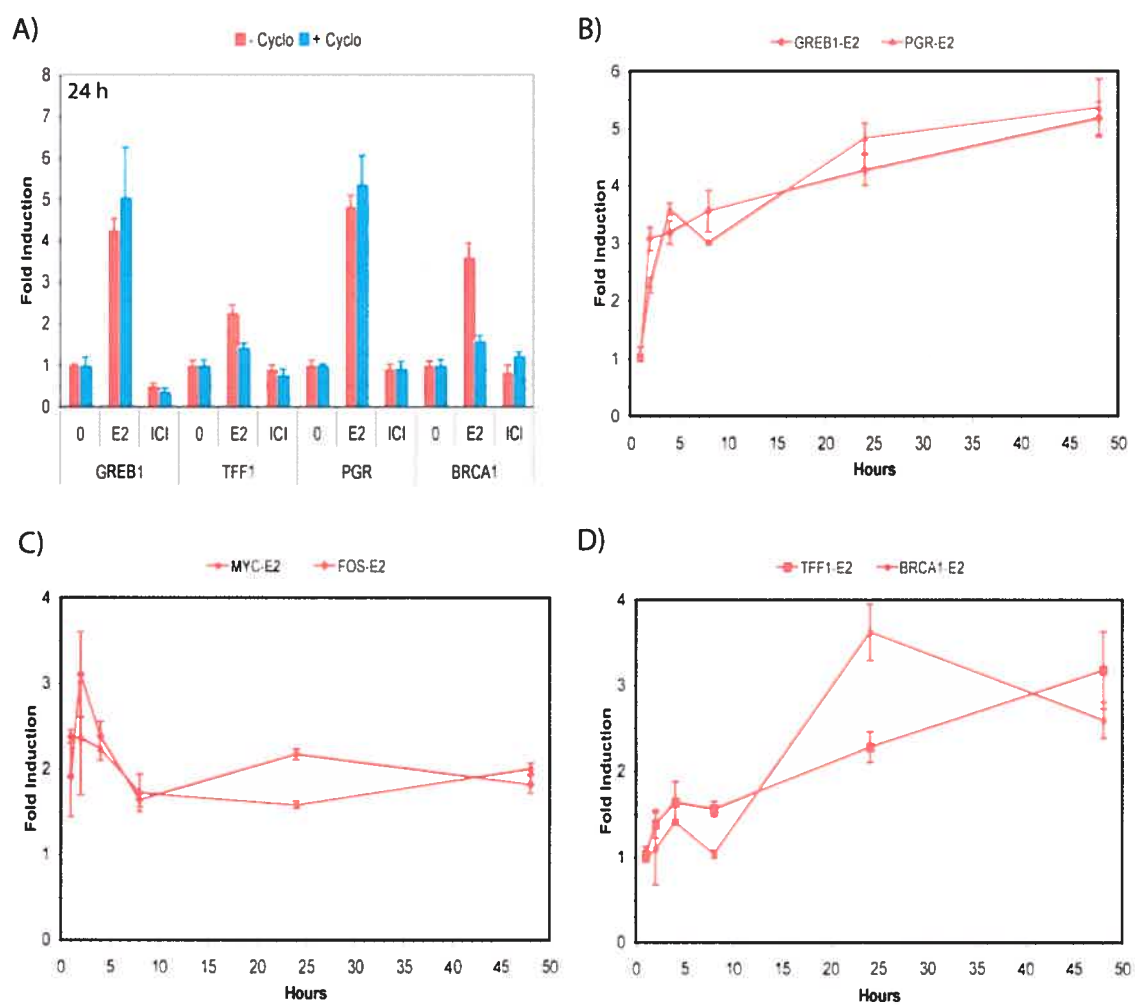
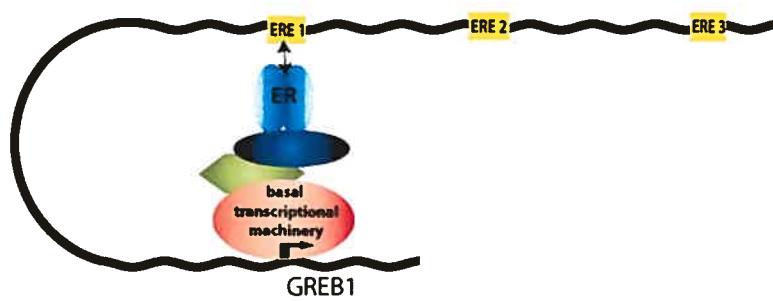
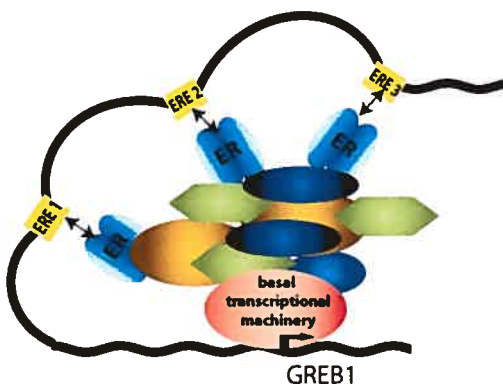


Figure Supplémentaire 3

without E2



with E2



Troisième partie

Chapitre 8 : Discussion et perspectives

Chapitre 8 : Discussion et perspectives

La signalisation découlant de la liaison des oestrogènes avec leurs récepteurs est diversifiée, spécifique à la fonction des tissus cibles et peut être altéré lors de l'apparition de pathologies. L'importance de comprendre les mécanismes de régulation de la prolifération cellulaire provenant des oestrogènes est primordiale dans la problématique de prévenir et/ou de développer de nouvelles avenues thérapeutiques pour le cancer du sein. Afin de comprendre et d'approfondir les causes d'une prolifération excessive lors de la formation de cancers mammaires, plusieurs façons d'aborder la problématique peuvent être mises en oeuvre. Lors de ma thèse, j'ai participé à la réalisation de différents projets ayant pour but général une meilleure compréhension des mécanismes de régulation de l'expression des gènes par les oestrogènes. Ces études ouvrent la porte à l'identification de gènes candidats impliqués dans la prolifération excessive des cellules de tumeurs mammaires.

Les études réalisées tout au long de ma thèse avaient essentiellement pour but de caractériser les mécanismes de la régulation des gènes par les oestrogènes et leur récepteur nucléaire ER α dans le tissu mammaire. Dans la première étude à laquelle j'ai participé, la spécificité et l'affinité de l'interaction récepteur-ADN a été abordée, permettant de mieux caractériser ce qui constitue un bon site de fixation pour ER et le mécanisme de liaison du récepteur. (annexe 1). Dans la deuxième étude, l'approche bio-informatique a permis l'élaboration d'une base de données de sites de fixation potentiels de haute affinité dans les génomes humains et de souris, apportant également une preuve de la fonctionnalité des *enhancers* distaux dans la régulation des gènes cibles des oestrogènes (chapitre 5). Dans la troisième étude, l'identification des cibles directes des oestrogènes a été réalisée en plus d'une recherche exhaustive des sites de fixation de facteurs de transcription pouvant contrôler la régulation de ces gènes, y compris des ERE (chapitre 6). Ceci a permis de démontrer un enrichissement en ERE dans les gènes cibles primaires des oestrogènes et une liaison démontrée du récepteur sur la plupart des ERE testés. Dans la dernière partie, une

étude approfondie a été réalisée afin de vérifier la fonctionnalité des *enhancers* distaux aussi bien que proximaux, démontrant que les mécanismes de régulation des gènes sont plus complexes qu'initialement prévus et pourraient impliquer des collaborations entre différents éléments répartis sur de grandes distances génomiques (chapitre 7).

Plusieurs concepts nouveaux se dégagent de ces différentes études en ce qui concerne la mécanistique de la régulation de la transcription par les oestrogènes. Certains faits ou concepts méritent que l'on s'y attarde plus en profondeur. Pour la première fois, nos études ont démontré que les ERE, à part un enrichissement très local autour du site d'initiation de la transcription, sont répartis uniformément dans l'ensemble du génome. Cette présence de fixation du récepteur dans des régions parfois éloignées des gènes nécessite d'approfondir la mécanistique de la transcription sur de longues distances pour mieux comprendre leur rôle dans la détermination des gènes directement régulés par ER α . Les directions futures de cette recherche portant sur la signalisation par les oestrogènes et l'analyse des mécanismes des effets médiés par les oestrogènes seront discutés.

8.1 Fixation à l'ADN de ER à l'échelle du génome

Les oestrogènes médient leurs effets par différentes voies : génomique classique, génomique alternative et non-génomique (comme décrit au chapitre 2). Même si la voie de signalisation classique semblait prépondérante pour la signalisation par les oestrogènes, peu de sites ERE validés *in vivo* étaient répertoriés dans la littérature (36). Plusieurs observations (gènes de la β -globine, récepteur des androgènes) (102,115,116,127) démontraient cependant que les facteurs de transcription pouvaient réguler la transcription sur de longues distances génomiques. La disponibilité des séquences complètes de plusieurs génomes a donc permis d'entreprendre des recherches bioinformatiques d'identification à grande échelle des ERE plutôt que seulement dans les quelques centaines

de paires de bases autour du site d'initiation de la transcription traditionnellement étudiée dans les études réalisées avec des gènes rapporteurs.

8.1.1 La prédiction de site de fixation de ER *in silico*

Deux voies principales pouvaient être entreprises pour l'identification bioinformatique d'EREs à l'échelle du génome : la recherche systématique des ERE et la recherche par matrice de poids. Dans la première étude (Chapitre 5), étant donné le petit nombre de ERE validés à cette époque, les matrices ont été écartées dû à leur biais possible. En effet, la pondération appliquée pour chaque position n'est valable que si le nombre de séquences utilisées pour bâtir la matrice est élevé et que chaque position du site de liaison possède une pondération représentant le plus fidèlement l'affinité du récepteur (128). Une recherche systématique des ERE de haute affinité a donc été retenue dans un premier temps. Pour ce faire, il a été nécessaire mieux définir ces sites de façon à rechercher dans le génome des séquences ayant un potentiel excellent d'être des ERE fonctionnels. L'étude réalisée au chapitre 5 a permis d'identifier plus de 70 000 ERE de haute affinité distribués dans le génome humain, de démontrer le caractère fonctionnel des ERE distaux en tant que sites de liaison pour le récepteur et d'établir l'importance de la structure de la chromatine sur la fixation des éléments de réponses.

Une observation tirée de cette étude était que la distribution des ERE dans le génome est uniforme, le nombre de ERE par chromosome étant proportionnel à la taille du chromosome (Chapitre 5, Fig 2). Lors d'analyse plus fine des données, on observe une surreprésentation des ERE (64%) autour de la région d'initiation de la transcription des gènes (-1000 pb à +2000 pb). Différentes études d'immunoprécipitation de la chromatine suivies d'hybridation sur une puce génomique (ChIP-CHIP) (103,104) démontrent que ces observations de concentration de sites de liaison (ou de sites liés) près des régions d'initiation de la transcription ne sont pas limitées au récepteur des oestrogènes.

Cette première étude a permis d'avoir une vue d'ensemble de la distribution des ERE dans les génomes. La restriction de se limiter aux séquences de hautes affinités et la

recherche systématique excluait potentiellement des sites de fixation importants dans la transmission des signaux oestrogéniques. À l'aide de notre première étude et de la littérature, nous avons bâti une matrice de poids représentant un grand nombre d'ERE fixés *in vivo* (Chapitre 6, Fig 6), ce qui nous a permis d'étendre notre recherche à des séquences plus distantes du consensus en limitant les biais de composition en nucléotides résultants d'un petit échantillon de séquences validées (129-131).

Cette nouvelle matrice (Chapitre 6 Fig 6), basée sur 76 ERE répertoriés et validés *in vivo*, a cependant aussi ses limites. Lorsque l'on compare la pondération dans la matrice pour certaines positions du ERE (particulièrement les positions 2,3,5) et les résultats des liaisons *in vitro* générées dans l'étude précédente (Chapitre 5 Fig 1), on peut facilement constater que certaines variations nuisibles à la liaison du récepteur ne sont pas parfaitement prises en compte. Ceci s'explique par le fait que le nombre et les types de séquences utilisées sont encore limités et ne représentent donc peut-être pas parfaitement l'affinité du récepteur pour l'ADN. Dans le futur, il serait possible de modifier la matrice par pondération selon les résultats d'affinité entre le récepteur et les ERE variants (Chapitre 5 Fig 1). Afin de voir si cette nouvelle matrice permet une détection améliorée des ERE selon l'affinité réelle du récepteur, il faudrait approfondir les études de liaison *in vitro* avec l'usage d'une plus grande combinaison de variations et comparer les prédictions de la matrice avec les résultats. Cependant, comme démontré dans l'annexe 1, les différentes positions dans l'ERE ne sont pas indépendantes donc il faudrait en tenir compte dans une modélisation future. Le changement d'un nucléotide dans la séquence peut influencer un autre nucléotide en venant perturber les interactions de ce dernier avec les acides aminés du récepteur. Ce type de relation peut être étudié par la construction de matrices de di-nucléotides et en comparant les fréquences de di-nucléotides prédites avec les modèles de liaisons du récepteur à chaque position du ERE. La définition d'un modèle informatique reflétant le plus fidèlement possible les variations d'affinité du récepteur est l'étape la plus importante dans ce type de recherche. À l'heure où la masse d'information engendrée par le séquençage des génomes et les coûts de la recherche biologique augmentent, des outils

informatiques fiables s'avèrent d'une grande importance afin de mieux interpréter les données d'expériences à grande échelle.

8.1.2 Les sites de liaison prédits sont-ils liés *in vivo*?

La recherche de sites de fixation ne se limite pas à les trouver, encore faut-il valider leur liaison *in vivo* et leur impact sur la régulation des gènes environnants. Dans ce but, nous avons jumelé nos études bio-informatiques à une étude à grande échelle identifiant les gènes régulés de façon primaire par les oestrogènes dans un modèle cellulaire dérivé d'un carcinome mammaire (Chapitre 6). Les ERE prédits à proximité de gènes cibles primaires représentent donc des candidats de choix pour valider leur liaison *in vivo* par une approche de ChIP. Simultanément à cette étude, un autre groupe a publié une étude de ChIP-CHIP de ER α sur toutes les régions non répétées du génome humain en présence d'oestrogènes dans les cellules MCF7 (132). Ces informations nous ont permis de comparer et de valider nos résultats de prédiction de sites (Chapitre 6). Nous avons recherché les ERE tels que décrit par notre matrice dans les régions liées par le récepteur dans le ChIP-CHIP (132). Ceci nous permet de corréler la présence d'un ERE potentiel et la fixation du récepteur *in vivo*. Nous restons conscients cependant que la technique de ChIP n'identifie pas avec précision le site fixé par le récepteur, qui se retrouve à l'intérieur de régions de chromatine liées de taille variant de 100 à plus de mille bp. De plus, cette technique ne discrimine pas entre une liaison directe du récepteur sur un ERE et les autres mécanismes alternatifs de liaison du récepteur (en tant que coactivateur ou sur un site de fixation hybride, Chapitre 2), et peut même générer des artéfacts dû au choix de l'anticorps. Cependant, la convergence des données de régulation de l'expression des gènes et la prédiction de sites de fixation et de ChIP constituent le meilleur indice possible du rôle de cet élément dans la liaison du récepteur et la régulation de la transcription, permettant de sélectionner cet élément pour des validations ultérieures.

L'absence de validation d'un ERE par ChIP-CHIP ne constitue pas non plus une preuve de son absence d'importance. Le fait d'avoir réalisé ces «validations» avec les

données d'un autre groupe nous limite dans l'interprétation car les conditions expérimentales utilisées entre nos deux groupes sont différentes (temps de traitement, concentration de sérum, concentration d'hormone). De plus, comme ce genre de technique est dépendant des réactifs utilisés (anticorps) et de l'exposition adéquate des épitopes reconnus par ces anticorps, on ne peut pas conclure que les sites prédits qui ne se retrouvent pas liés dans l'étude de CHIP-CHIP sont faux simplement sur cette comparaison. De plus, un site qui n'est pas fixé par le récepteur dans un type cellulaire peut être fonctionnel dans un autre modèle cellulaire. Dans un monde idéal, il faudrait générer plusieurs de ces études à grande échelle dans les différentes lignées cibles des oestrogènes et sous différentes conditions (temps, concentration d'hormone et de sérum), avec différents anticorps. Ainsi, nous aurions un tableau plus représentatif des sites ERE prédits qui peuvent être fonctionnels *in vivo*. Ces connaissances permettraient ensuite d'analyser quels facteurs déterminent l'accessibilité d'un ERE au récepteur, incluant probablement la structure de la chromatine et la présence de différents partenaires du récepteur dans la cellule.

8.2 Gènes directement régulés par ER α

De façon à comprendre le rôle de la régulation transcriptionnelle médiée par ER α dans la médiation du signal prolifératif des oestrogènes, il est important de caractériser de façon exhaustive les gènes cibles de ER α . Les études réalisées sur l'expression des gènes dans des tumeurs mammaires et/ou provenant de prélèvements de tissus sains, nous apportent un tableau des gènes exprimés globalement, sans discriminer le rôle des facteurs de croissance et des hormones diverses présentes dans la circulation sanguine. Une façon de parvenir à établir le rôle des œstrogènes est d'utiliser des modèles de cellules en culture. Des cellules primaires normales, immortalisées et/ou des cellules dérivées de carcinomes peuvent être utilisées. La difficulté d'utiliser des cellules primaires, outre l'accessibilité restreinte et réglementée, est d'obtenir une population composée d'un seul type de cellules et de conserver l'expression du récepteur ER α . En effet, une perte de l'expression du

récepteur lors de la culture primaire des cellules épithéliales a été observée. Les lignées cellulaires dérivées de carcinomes mammaires conservent par contre une expression forte du récepteur permettant l'étude de la prolifération stimulée par les oestrogènes dans les tumeurs mammaires.

8.2.1 ER α médiateur de la signalisation oestrogénique pour tous les gènes régulés?

Dans le chapitre 6, nous avons généré une étude d'expression des gènes régulés en présence d'oestrogène dans une lignée cellulaire dérivée de carcinome mammaire, les MCF-7. Le choix de nos conditions expérimentales s'est fait de manière à utiliser des conditions propices à la prolifération cellulaire et à l'identification des cibles primaires des oestrogènes (Chapitre 6 Fig1 & Fig S2). Lors de l'analyse, nous avons classé les gènes selon leur régulation oestrogénique (*Up* ou *Down*) ainsi que l'influence de la cycloheximide sur leur régulation. La proportion des gènes régulés positivement vs négativement en présence d'oestrogènes est similaire (1214 vs 955). Par contre une plus grande différence est observée lorsque l'on conserve seulement les gènes primaires (544 *Up* vs 236 *Down*). Ceci laisse présager que les facteurs impliqués dans la signalisation jouent un rôle dans la stimulation de la transcription plutôt que dans sa répression. En recherchant par analyse bio-informatique les sites de fixation potentiels de ER autour de ces gènes, on retrouve un enrichissement marqué des ERE sur les promoteurs des gènes primaires *Up* régulés. Par contre, un appauvrissement de ERE est observé à proximité des gènes primaires *Down* régulés. Nous avons également cherché à identifier les autres facteurs de transcription qui pourraient être des partenaires potentiels à ER α pour la transcription des gènes dans nos différentes catégories par analyse de l'enrichissement de leur site de liaison dans les séquences flanquantes de gènes. Quelques sites de facteurs de transcription sont ressortis enrichis dans les gènes réprimés (*Down*), cependant nous avons dû rechercher dans le promoteur proximal (-/+ 2,5 Kpb) et nous concentrer sur les 100 meilleurs gènes régulés pour obtenir une signification statistique (Chapitre 6, Table S1, S2). L'inconvénient majeur de ce genre de recherche bio-informatique est la définition des sites de fixations de facteurs

de transcription. Comme discuté précédemment, les recherches bio-informatiques par matrices sont très utiles si les matrices sont réellement représentatives de l'affinité du facteur de transcription pour l'ADN. Nos recherches pour identifier d'autres facteurs de transcription impliqués dans l'expression de nos gènes *Down* régulés ont donc été limitées par les différentes matrices disponibles. Un autre problème dans l'analyse des mécanismes de régulation est la taille restreinte en nombre de gènes de certaines catégories, rendant impossible d'effectuer des statistiques adéquates et donc de générer des conclusions fiables sur les médiateurs potentiels de cette régulation négative.

Des études approfondies seront nécessaires pour comprendre quel est le mécanisme de régulation de la transcription des gènes réprimés. Comme peu de sites ERE ont été retrouvés, ceci suggère qu'une bonne partie de ces gènes pourraient être régulés par des voies de signalisation non classiques. Différentes études biologiques pourraient être réalisées pour découvrir les mécanismes de signalisation impliqués pour ces gènes. Par exemple, l'abolition de ER α dans les cellules MCF7 par shRNA ou « ARN tige-boucle » stablement exprimé permettrait de déterminer par analyse de micropuce quels sont les gènes qui sont encore réprimés par les oestrogènes malgré la perte du récepteur. Les signalisations non classiques des oestrogènes pourraient aussi être évaluées par le même genre de technique, par exemple l'implication des facteurs JUN/FOS (AP1), la signalisation membranaire de ER α , le récepteur membranaire GPR30, et les autres facteurs de transcription que nous avons identifiés dans la recherche bio-informatique.

8.2.2 Les rôles fonctionnels des gènes régulés directement par ER

L'étude réalisée dans le chapitre 6 a permis d'identifier des cibles primaires des oestrogènes par leur régulation en présence de cycloheximide et par la possibilité de la liaison du récepteur dans leurs séquences flanquantes. Il est probable que plusieurs de ces gènes médient l'effet prolifératif des oestrogènes, comme déjà démontré pour les gènes CCND1, MYC, FOS etc... Des cibles primaires des oestrogènes impliquées dans leur effet prolifératif et qui seraient spécifiques aux cellules mammaires constitueraient de bonnes

cibles thérapeutiques. L'identification générale du rôle des cibles primaires des oestrogènes pourrait aussi permettre de révéler de nouvelles fonctions des oestrogènes dans les cellules mammaires. Un objectif important sera donc d'analyser la fonction de l'ensemble des gènes cibles primaires des oestrogènes.

Lors d'analyses effectuées avec un logiciel d'analyse de fonction des gènes (FATIGO), nous avons observé peu de données claires sur l'enrichissement d'une classe particulière de gènes dans nos cibles primaires. La limitation de la recherche avec ce programme est le nombre de gènes qu'il est possible d'entrer comme données comparatives, ce qui rend difficile l'étude d'enrichissement fonctionnel face à l'ensemble des gènes présents sur la puce de départ. Bien que les outils informatiques disponibles présentement nous limitent sur l'étude d'enrichissements fonctionnels, l'utilisation de programmes développés spécifiquement pour l'étude des réseaux de régulation cellulaire (ex : IPA 5.0) pourrait nous donner une vision globale des types de réseaux de gènes retrouvés dans nos différentes catégories (*Up*-primaire, *Down*-primaire ou secondaire). Ceci nous permettrait d'établir des corrélations fonctionnelles entre les différentes catégories de gènes et d'avoir une indication précise du rôle des gènes cibles primaires dans notre lignée mammaire. Afin de palier aux différentes limitations des outils informatiques disponibles, des recherches systématiques sur la fonction des gènes, bien que longues, peuvent être réalisées afin d'atteindre le but qui est l'identification des gènes primaires de ER impliqués dans la prolifération cellulaire.

La sur-expression et/ou l'abolition de l'expression d'un gène candidat pourra être réalisée dans des cellules MCF-7, afin de voir l'impact de ce changement sur la progression du cycle cellulaire, la prolifération cellulaire et la migration cellulaire. L'avantage d'utiliser à la fois l'abolissement et la sur-expression d'un gène est d'éviter les artéfacts des différentes méthodes. La sur-expression peut finir par saturer la machinerie cellulaire et ainsi masquer un phénotype. L'abolition par des outils comme les shRNA peut causer des phénotypes qui ne sont pas spécifiques au gène étudié mais dus plutôt à un effet sur une autre cible. Des études d'abolition ou de sur-expression dans la glande mammaire de

modèles de souris permettront par la suite de déterminer l'impact sur le développement de tumeurs de ces gènes candidats. La poursuite de cette étude en ciblant des gènes impliqués dans la prolifération permettra de valider le rôle crucial des gènes primaires dans la prolifération cellulaire médiée par les oestrogènes. Un très bon exemple de gène candidat pertinent à un approfondissement fonctionnel dans les tumeurs mammaires est GREB1. Bien que la fonction précise de GREB1 soit encore inconnue, il a été démontré qu'il est non seulement un gène cible primaire de ER α (Chapitre 7 fig 1 et S2) mais est aussi exprimé plus fortement dans les tumeurs mammaires ER positives (Chapitre 7, fig S1). De plus, une approche de suppression de l'expression de ce gène par siRNA a suggéré un rôle dans les effets prolifératifs de E2 (133). L'effet de la modulation de l'expression de GREB1 sur la tumorigénèse dans des modèles animaux reste à examiner.

8.3 Mécanistique de la transcription sur de longues distances

Plusieurs études ont démontré que des sites de liaison distaux de différents facteurs de transcription peuvent être fonctionnels, à la fois pour rendre la chromatine plus accessible aux autres protéines, et en participant directement à la concentration des protéines nécessaires à la régulation transcriptionnelle du gène (114,116,117,134-137). Dans le système des gènes de la globine, il est bien établi que les régions LCR lointaines (environ 50Kpb) participent à la régulation de la transcription des gènes par la formation de boucle de chromatine. Cependant, peu d'études de ce genre (105,138) ont examiné si ce type de régulation est applicable aux récepteurs nucléaires, et en particulier si ER α peut former des boucles de chromatine servant à concentrer les récepteurs et leurs cofacteurs sur un promoteur donné, participant ainsi à la transcription.

8.3.1 Les ERE distaux sont impliqués dans la formation de boucle de chromatine.

Le gène GREB1, qui est une cible primaire des oestrogènes et semble impliqué dans la prolifération cellulaire médiée par les oestrogènes (133), est un bon modèle d'étude pour la mécanistique de la transcription sur de longue distance. Il est exprimé dans les tumeurs mammaires et représente un cas de régulation possédant trois ERE consensus distribués sur 21 Kpb. Dans les conditions expérimentales choisies, les trois ERE sont fixés par ER α avec une cinétique d'association comparable, et la chromatine environnante est acétylée (chromatine permissive) (Chapitre 7, Fig 2). Des études de formation de boucles de chromatine (technique 3C) ont déterminé que les éléments s'associent à la région promotrice et s'associent entre eux (Chapitre 7, Fig 3). Une différence majeure dans notre système avec ce qui a été précédemment démontré pour le gène de la β -globine est que la formation des boucles est affectée par l'hormone ou le composé lié à ER, donc inductible selon les ligands présents. Lors de l'ajout d'agonistes de la transcription (œstrogènes), il y a formation de boucle permettant aux éléments distaux de venir participer à la transcription du gène. Par opposition, lors de l'utilisation d'anti-œstrogène totaux il n'y a aucune formation de boucle entre ces éléments et le site d'initiation de la transcription. Cette observation est en accord avec le fait que ces composés antagonistes ne permettent pas à ER de transcrire GREB1. Finalement, seulement l'élément le plus proche du site d'initiation de la transcription forme une interaction forte en absence d'hormone, suggérant ainsi que la transcription basale du gène implique peu les deux éléments distaux.

De plus, nos résultats démontrent clairement pour la première fois que ER α participe à la régulation de la transcription sur de longues distances en formant des boucles de chromatines entre plusieurs *enhancers*. Ce mécanisme pourrait être généralisable, car les études de ChIP-CHIP indiquent que plusieurs régions de fixation sont fréquemment retrouvées à proximité des gènes cibles des oestrogènes.

Certains éléments d'information supplémentaires pourraient être obtenus dans des études subséquentes sur les mécanismes de formation de ces boucles de chromatine. Premièrement, il serait souhaitable de démontrer de façon directe que ER α est impliqué dans la formation des boucles par des études de ChIP-3C (138). Deuxièmement, les essais d'immunoprécipitations de la chromatine ont démontré le recrutement des mêmes cofacteurs sur les trois ERE. Ce qui suggère que les boucles sont formées par un pont de protéines, probablement induits par le recrutement du récepteur en présence d'œstrogène sur les éléments distaux. Des études de ChIP permettront de déterminer quels sont les coactivateurs responsables de la transcription et leur implication dans la formation des boucles pourra être déterminée par ChIP-3C. Le rôle individuel des EREs dans la régulation transcriptionnelle et la formation de boucles pourront être examinés par l'abolition des sites *enhancers* par différentes techniques (inactivation de la chromatine, délétion du site) à l'aide de protéines à doigt de zinc (139). Ces études permettront d'approfondir comment ER α peut médier une transcription de gènes via une fixation sur des sites lointains. Les différentes études (Chapitre 3) sur l'implication et le rôle des sites de fixation lointains des facteurs de transcription ont permis de développer plusieurs hypothèses sur le rôle potentiel d'une boucle de chromatine : la formation d'une boucle par des liens protéiques entre des facteurs de transcriptions et des cofacteurs, simplement un mécanisme de collisions des chromosomes favorables à la transcription, la formation de boucles facilitées par le cheminement de la polymérase (117)? L'approfondissement de la mécanistique des boucles de chromatines médiées par les ER pourra permettre de déterminer le rôle de ces boucles dans la transcription oestrogénique.

Les boucles que nous avons observées se situaient dans le temps après 45 min de stimulation des cellules MCF7 par les oestrogènes. L'analyse de la cinétique de recrutement de ER α en présence d'oestrogènes a démontré que le recrutement le plus important du récepteur sur les ERE se fait à 45 min (Chapitre 7 Fig 2). Cependant, GREB1 est un gène qui possède une stimulation soutenue dans le temps, dépassant 72H. Est-ce que les boucles sont toujours présentes dans le temps, ou bien disparaissent-elles après que la transcription soit amorcée? Leur présence est-elle due seulement à l'initiation de la

transcription lors de l'arrivée de l'hormone? Pour répondre à ces différentes questions, nous devrions étudier le recrutement des protéines sur les ERE et les sites d'initiation de la transcription, l'état de la chromatine et la formation de boucles dans le temps. Le même genre de techniques précédemment utilisées (ChIP et 3C) pourront être mise en oeuvre pour approfondir la dynamique de formation de ces boucles.

La formation des boucles est majoritairement inductible en présence d'agonistes et abolit en présence d'un antagoniste total. Il a été démontré dans le chapitre 7 que la transcription basale du GREB1 est diminuée par ICI 182 780 (fig S2). Comment la présence de l'antagoniste total empêche-t-elle la formation des boucles qui sont présentes à l'état basal (Chapitre 7, fig 3)? Est-ce que la liaison de cette molécule antagoniste provoque le recrutement de corépresseurs qui empêcheraient la formation des boucles de chromatines (140,141)? Est-ce que la liaison de ICI entraîne la dégradation et/ou la séquestration (48,142-145) de $ER\alpha$, prévenant la transcription basale? Il serait bon de déterminer par des études classiques de ChIP quelles sont les protéines associées en présence de ICI avec $ER\alpha$ sur le gène GREB1. Selon les résultats obtenus, on pourrait identifier le mécanisme potentiel qui régit l'abolition des boucles de chromatines qui sont formées lors de la transcription basale de GREB1.

Finalement, lorsque le promoteur du gène GREB1 est analysé sur NCBI (Aceview), plusieurs isoformes sont décrits. Les sites d'initiation de la transcription de certains isoformes sont prédits beaucoup plus en aval que celui de la forme que nous avons étudiée (l'isoforme «A», le plus en 5'). Nos données démontrent que ces autres isoformes sont régulés par les oestrogènes avec des cinétiques comparables à celle de la forme «A» dans les cellules MCF7. Il serait donc intéressant de poursuivre l'étude de la mécanistique de la transcription de cette région génomique en incluant les autres isoformes de GREB1. Ceci permettrait de déterminer si les ERE (qui se retrouvent pour certains isoformes à -50Kpb) participent directement à la transcription de ces autres isoformes. Bien que les isoformes soient transcrits de façon similaire dans les cellules MCF7, on peut imaginer une transcription différentielle des isoformes dans un autre type cellulaire. Finalement en amont

de l'isoforme A, à une distance équivalente à certains isoformes, on retrouve sur le brin inverse le gène E2F6. Est-il lui aussi influencé par la présence de ces ERE, surtout que nous avons démontré sa régulation dans le chapitre 6. La formation de boucles sur les autres sites d'initiation de GREB1 et sur E2F6 ouvriraient donc la porte à une mécanistique complexe de la transcription du locus.

8.3.2 Fonctionnalité des boucles de chromatine dans les glandes mammaires

Un aspect qui serait intéressant pour la compréhension du rôle fonctionnel de la formation de ces boucles serait d'étudier ce phénomène *in vivo*. Quelle est la place de ces boucles de chromatines dans la transcription, sont-elles présentes dans l'expression régulière de GREB1 (glande mammaire normale), apparaissent-elles lors de la maturation de la glande mammaire, ou bien lors de la tumorigenèse? Est-ce qu'elles sont conservées entre les espèces ou bien est-ce que l'utilisation de ces ERE lointains est un mécanisme de développement tumorale observable seulement dans nos cellules MCF7? Pour répondre à ces questions, la régulation de la transcription devrait être analysée dans un premier temps dans des cellules mammaires de souris de façon à déterminer la conservation de ce phénomène au travers des espèces. Comme une partie des ERE est conservé chez la souris (Chapitre 5), on devrait s'attendre à une conservation de ce mécanisme de régulation de la transcription dans des lignées cellulaires de souris. Par la suite, si la régulation du gène par les oestrogènes est conservée, il serait intéressant d'étudier les mécanismes de transcription dans les glandes mammaires de souris au courant de leur développement, de leur maturation ou lors du développement tumorale (souris immatures, lors de la puberté, ou modèle tumorale de glandes mammaires). En plus d'apporter une meilleure compréhension de la transcription, ces études permettraient peut-être de mettre en lumière des indices sur la fonction des boucles. Par exemple, si elles apparaissaient seulement lors du développement tumoral, on pourrait conclure que la régulation de la transcription sur de longue distance n'est pas utile pour avoir les niveaux requis de GREB1 dans une cellule saine. Cependant, si elles sont observées lors de la maturation des glandes mammaires, une augmentation

significative de la transcription de GREB1 doit avoir lieu pour promouvoir la maturation et la prolifération des tissus mammaires.

8.4 L'intégration des réseaux de régulation de la transcription

Bien que le recrutement du ER sur son ERE soit suffisant à l'activation transcriptionnelle de promoteurs de base, les séquences régulatrices des gènes contiennent souvent de multiples sites de facteurs de transcription, soulevant la question de savoir si certains de ces facteurs sont des partenaires privilégiés pour la régulation transcriptionnelle par ER. De plus, certains gènes primaires sont eux-mêmes des facteurs de transcription (JUN, FOS, MYC), et permettent une amplification du signal des œstrogènes à des cibles secondaires. Des boucles de régulation positives ou négatives peuvent ainsi être créées lorsque des facteurs de transcription ou cofacteurs sont à la fois impliquer dans la régulation primaire de la transcription par ER α et sont également des cibles de ce récepteur.

8.4.1 Coopérativité de ER α avec d'autres facteurs de transcription

La caractérisation des régions flanquantes des gènes cibles primaires des oestrogènes indique que certains de ces gènes (*Up* et particulièrement *Down*) ne possèdent pas d'ERE à une distance de 25 Kpb de leur site d'initiation de la transcription. Bien que l'existence d'ERE plus distaux ne puisse être exclu, ces observations suggèrent l'utilisation de voies de signalisations oestrogéniques non classiques, qui peuvent inclure le recrutement de récepteurs par l'intermédiaire d'autres facteurs de transcription ou l'activation d'autres facteurs de transcription. Des recherches bioinformatiques ont conduit à l'identification de sites de fixation de facteurs de transcription dans les séquences flanquantes de ces gènes (chapitre 6). Il serait intéressant de compléter cette étude par une analyse de la sur-représentation potentielle de ces mêmes sites dans les données de ChIP-CHIP. Ceci indiquerait leur rôle dans le recrutement indirect de ER α à l'ADN (mécanisme de

tethering). Cette hypothèse pourrait être confirmée par une approche de ChIP-re-ChIP démontrant le co-recrutement de ER α avec le facteur de transcription partenaire au site identifié. Cependant, un enrichissement dans les régions flanquantes de gènes cibles mais non dans les régions de ChIP pourrait s'expliquer par l'activation du facteur de transcription par des voies non-génomiques. Une vérification de l'implication de facteurs de transcription partenaires dans la régulation transcriptionnelle par ER α passerait par l'étude de l'influence des oestrogènes sur l'expression dirigée par un vecteur rapporteur contenant le site de fixation du facteur de transcription.

Bien que plusieurs exemples de facteurs de transcription partenaires de ER α soient connus (AP1, Sp1, NF κ B etc...), il n'est pas clair quel est le degré de généralité de ces mécanismes de régulation. La plupart des facteurs de transcription interagissent probablement de façon indirecte par l'établissement de ponts protéiques impliquant le médiateur ou la machinerie transcriptionnelle de base. Ces interactions sont susceptibles de médier la coopération transcriptionnelle entre plusieurs sites liés par des facteurs de transcription différents pour les séquences de régulation d'un même gène. Cependant, la force de cette coopérativité pourrait dépendre de celle des interactions entre facteurs de transcription, qui peuvent être directes ou médiées par une ou plusieurs protéines.

Afin d'étudier les partenaires potentiels de coopérativité transcriptionnelle de ER, une approche possible serait d'utiliser des études de boucles de chromatine à grande échelle (146,147) qui permettraient de caractériser les interactions entre un ERE donné (par exemple un des ERE de GREB1) et d'autres sites dans la même région chromosomique (*cis* régulation) et/ou à plus grandes distances voir entre chromosomes (*trans* régulation). Par la suite, des analyses bio-informatiques pourraient nous indiquer les partenaires potentiels en recherchant dans ces régions bouclées des sites de fixation de différents facteurs de transcription. L'information obtenue donnerait des pistes sur la diversité des partenaires possibles. Bien que l'utilisation de ces deux techniques soit coûteuse, elle permettrait de générer des informations précieuses à grande échelle sur les voies de signalisation

convergentes en présence d'oestrogènes. De plus, nous pourrions identifier un rôle potentiel de ces EREs dans la régulation de promoteurs très lointains

Certaines hypothèses intéressantes sur l'organisation spatiale de la chromatine dans le noyau, ont permis d'élaborer des théories reliant les boucles de chromatine et la convergence des réseaux de transcription. Il semble exister une formation de régions de transcription dans le noyau via de larges boucles radiales (101, 100). Ces régions souvent associées à la matrice nucléaire se comporteraient comme des petites usines de transcription qui pourraient transcrire plusieurs gènes reliés entre eux sous un traitement particulier dans un même espace tridimensionnel. On pourrait donc penser que les interactions entre les différents facteurs de transcriptions permettraient la formation de ces régions «particulières» de transcription.

8.4.2 L'amplification du signal oestrogénique

L'autre niveau de réseau de régulation dérivé de nos études consiste en la mise en relation des cibles primaires et secondaires des oestrogènes. Les facteurs de transcription régulés de façon primaire (FOS, JUN, MYC) peuvent à leur tour médier l'expression de leurs gènes cibles respectifs (cibles secondaires). Mieux comprendre le rôle des cibles primaires dans la propagation des effets oestrogéniques est important pour une vision globale des réseaux de signalisation dans les tissus mammaires.

Des études d'expression des gènes cibles des oestrogènes dans des cellules MCF7 exprimant de façon stable des shRNA pour nos facteurs de transcription régulés par les oestrogènes de façon primaire permettraient de vérifier leur implication dans la régulation de gènes cibles. La régulation secondaire spécifique à un sous-réseau de régulation par un facteur de transcription donné serait ainsi perdue, ce qui permettrait d'identifier par leur absence les réseaux de régulation secondaires. Les observations réalisées dans le Chapitre 6 suggèrent que MYC ne serait pas un bon candidat pour ce genre d'étude (du moins dans un modèle de régulation de 24 hrs) puisque peu de gènes cibles de MYC sont enrichis dans nos données. Par contre, il se peut que les fenêtres dans lesquelles la sur-représentation des

sites myc a été étudiée ne soient pas adaptées. En effet, les sites de fixation de myc semblent se retrouver de manière prédominante dans les îlots CpG. Une recherche de sur-représentation focalisée sur ces îlots pourrait être plus concluante. D'un autre côté, E2F semble un facteur de transcription impliqué dans la propagation du signal estrogénique car nous avons trouvé une sur-représentation claire des sites de fixation de ces facteurs (Chapitre 6, Fig 9).

Une approche complémentaire pour étudier le rôle de facteurs de transcription dans la propagation des effets des oestrogènes impliquait la réalisation d'une expérience ChIP-CHIP dirigée contre ces facteurs en présence d'oestrogènes. Il suffirait de comparer les gènes fixés par ces facteurs avec nos gènes cibles secondaires pour identifier un réseau potentiel de d'aplification du signal.

Une étude récente démontre un bel exemple de ces réseaux de régulation et d'amplification du signal entre un gène direct, GATA3, qui est un facteur de transcription, et ER α (148). GATA3 est régulé par la liaison directe de ER α sur son promoteur, par la suite GATA3 va à son tour médier ses effets dans la cellule en ciblant des gènes dont ER α lui-même. Ceci confirme l'importance d'aller vers des études d'intégration du signal oestrogénique dans les cellules mammaires pour comprendre les mécanismes de la prolifération.

8.4.3 Etablissement de boucles de rétroaction

Le gène TFF1 est un bon exemple de régulation complexe impliquant des partenaires aux récepteurs des oestrogènes. ER α est connu pour être recruté de façon dynamique sur le promoteur et pour médier son expression par les oestrogènes. Cependant, comme nous l'avons observé dans le chapitre 7, la transcription de TFF1 est grandement affectée par la cycloheximide. Une partie de l'explication vient du fait que la famille des facteurs de transcription AP- peut elle aussi médier la transcription de ce gène et que

l'expression de JUN et FOS (facteurs se liant au site AP-1) est elle-même régulé par les oestrogènes (149).

Un autre exemple de boucle de rétroaction est la collaboration entre FOXA1 et ER α . Il a été démontré que la présence de FOXA1 à proximité de site de fixation de ER α était requise pour la fixation de ER α sur certains promoteurs (138,150). Cette relation qui est séquentielle et mutuellement exclusive (les deux facteurs ne sont pas présents au même moment sur l'ADN), démontre une collaboration de facteurs de transcription participant à l'expression d'un gène cible en présence d'oestrogènes.

En conclusion, l'identification de partenaires de coopérativité transcriptionnelle, la compréhension des liens qui les unient et l'étude de mécanismes de transcription croisée, permettra de broser un tableau plus large des réseaux de régulations en présence d'oestrogènes.

8.5 Conclusion

Les oestrogènes jouent un rôle dans la régulation de la prolifération des cellules épithéliales mammaires normales et cancéreuses. Le travail réalisé ici a permis d'identifier les gènes primaires des oestrogènes, de préciser les mécanismes de régulation classique de la transcription et d'esquisser les réseaux de régulation transcriptionnelle impliquant le ER α . Ces études permettent aussi d'ouvrir la porte à des recherches sur des candidats potentiels dans la prolifération des tumeurs mammaires ER positives. L'abolition systématique et/ou la surexpression de ces gènes pourrait être une avenue possible pour identifier de bons candidats, tout comme un approfondissement des fonctions et des patrons d'expression de ces différents gènes dans les tumeurs mammaires. Des études génomiques, à l'aide de récepteur chimérique (139,151,152) pourraient aussi être intéressantes à développer pour moduler l'expression de ces gènes

En plus de concentrer les directions futures sur les gènes primaires des oestrogènes participant au développement tumoral, il serait pertinent d'avoir en parallèle une vision globale de la signalisation des oestrogènes. Une bonne façon globale d'y parvenir, sans contrainte du choix de la voie de signalisation des oestrogènes, est l'établissement des réseaux de gènes dont la transcription est reliée en présence d'oestrogènes, selon les différents tissus. L'amélioration à grande échelle de la technique de boucles de chromatine lors d'une stimulation avec des oestrogènes permettrait de dégager des ensembles de gènes interagissant à un instant « t », dans un tissu particulier.

Plusieurs effets physiologiques des oestrogènes ont été présentés dans le début de cette thèse, dans le futur, avec le raffinement des différentes technologies d'études à grande échelle, il serait intéressant de revenir à des modèles plus globale. L'étude de cellules en boîte est pratique pour déterminer certains phénomènes précis, mais la compréhension complète des effets prolifératifs des oestrogènes passera à la fois par l'étude des oestrogènes dans différents tissus et par l'étude de ces tissus intégrés dans un organisme entier avec tous les stimuli auquel le corps est exposé.

Bibliographie

1. Robert, P., Rey-Debove, J. and Rey, A. (2007) In Robert, D. L. (ed.), Paris.
2. Meinhardt, U. and Mullis, P.E. (2002) The essential role of the aromatase/p450arom. *Semin Reprod Med*, **20**, 277-284.
3. Meinhardt, U. and Mullis, P.E. (2002) The aromatase cytochrome P-450 and its clinical impact. *Horm Res*, **57**, 145-152.
4. Simpson, E., Rubin, G., Clyne, C., Robertson, K., O'Donnell, L., Jones, M. and Davis, S. (2000) The role of local estrogen biosynthesis in males and females. *Trends Endocrinol Metab*, **11**, 184-188.
5. Poutanen, M., Isomaa, V., Peltoketo, H. and Vihko, R. (1995) Regulation of oestrogen action: role of 17 beta-hydroxysteroid dehydrogenases. *Ann Med*, **27**, 675-682.
6. Greenspan FS, S.G. (1997) *Basic & Clinical Endocrinology*. Fifth edition ed. Appleton & Lange, Stamford.
7. Nilsson, S., Makela, S., Treuter, E., Tujague, M., Thomsen, J., Andersson, G., Enmark, E., Pettersson, K., Warner, M. and Gustafsson, J.A. (2001) Mechanisms of estrogen action. *Physiol Rev*, **81**, 1535-1565.
8. Couse, J.F. and Korach, K.S. (1999) Estrogen receptor null mice: what have we learned and where will they lead us? *Endocr Rev*, **20**, 358-417.
9. Ciocca, D.R. and Roig, L.M. (1995) Estrogen receptors in human nontarget tissues: biological and clinical implications. *Endocr Rev*, **16**, 35-62.
10. Hess, R.A., Bunick, D., Lee, K.H., Bahr, J., Taylor, J.A., Korach, K.S. and Lubahn, D.B. (1997) A role for oestrogens in the male reproductive system. *Nature*, **390**, 509-512.
11. Bilinska, B., Wiszniewska, B., Kosiniak-Kamysz, K., Kotula-Balak, M., Gancarczyk, M., Hejmej, A., Sadowska, J., Marchlewicz, M., Kolasa, A. and Wenda-Rozewicka, L. (2006) Hormonal status of male reproductive system:

- androgens and estrogens in the testis and epididymis. In vivo and in vitro approaches. *Reprod Biol*, **6 Suppl 1**, 43-58.
12. Carreau, S., Lambard, S., Delalande, C., Denis-Galeraud, I., Bilinska, B. and Bourguiba, S. (2003) Aromatase expression and role of estrogens in male gonad : a review. *Reprod Biol Endocrinol*, **1**, 35.
 13. Imagawa, W., Yang, J., Guzman, R. and Nandi, S. (1994) *Control of mammary development*. Raven Press, New york.
 14. Russo, J., Ao, X., Grill, C. and Russo, I.H. (1999) Pattern of distribution of cells positive for estrogen receptor alpha and progesterone receptor in relation to proliferating cells in the mammary gland. *Breast cancer research and treatment*, **53**, 217-227.
 15. Saji, S., Jensen, E.V., Nilsson, S., Rylander, T., Warner, M. and Gustafsson, J.A. (2000) Estrogen receptors alpha and beta in the rodent mammary gland. *Proc Natl Acad Sci U S A*, **97**, 337-342.
 16. Hoskins, K. and Weber, B.L. (1994) The biology of breast cancer. *Current opinion in oncology*, **6**, 554-559.
 17. Wiesen, J.F., Young, P., Werb, Z. and Cunha, G.R. (1999) Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development. *Development (Cambridge, England)*, **126**, 335-344.
 18. Xie, W., Paterson, A.J., Chin, E., Nabell, L.M. and Kudlow, J.E. (1997) Targeted expression of a dominant negative epidermal growth factor receptor in the mammary gland of transgenic mice inhibits pubertal mammary duct development. *Mol Endocrinol*, **11**, 1766-1781.
 19. Ankrapp, D.P., Bennett, J.M. and Haslam, S.Z. (1998) Role of epidermal growth factor in the acquisition of ovarian steroid hormone responsiveness in the normal mouse mammary gland. *J Cell Physiol*, **174**, 251-260.
 20. Turner, R.T., Riggs, B.L. and Spelsberg, T.C. (1994) Skeletal effects of estrogen. *Endocr Rev*, **15**, 275-300.

21. Tsang, K.L., Jiang, H., Ramsden, D.B. and Ho, S.L. (2001) The use of estrogen in the treatment of Parkinson's disease. *Parkinsonism Relat Disord*, **8**, 133-137.
22. Singh, M., Dykens, J.A. and Simpkins, J.W. (2006) Novel mechanisms for estrogen-induced neuroprotection. *Experimental biology and medicine (Maywood, N.J)*, **231**, 514-521.
23. Kuiper, G.G., Shughrue, P.J., Merchenthaler, I. and Gustafsson, J.A. (1998) The estrogen receptor beta subtype: a novel mediator of estrogen action in neuroendocrine systems. *Front Neuroendocrinol*, **19**, 253-286.
24. Farhat, M.Y., Lavigne, M.C. and Ramwell, P.W. (1996) The vascular protective effects of estrogen. *Faseb J*, **10**, 615-624.
25. Barrett-Connor, E. and Grady, D. (1998) Hormone replacement therapy, heart disease, and other considerations. *Annual review of public health*, **19**, 55-72.
26. Nathan, L. and Chaudhuri, G. (1997) Estrogens and atherosclerosis. *Annual review of pharmacology and toxicology*, **37**, 477-515.
27. Lundeen, S.G., Carver, J.M., McKean, M.L. and Winneker, R.C. (1997) Characterization of the ovariectomized rat model for the evaluation of estrogen effects on plasma cholesterol levels. *Endocrinology*, **138**, 1552-1558.
28. Deroo, B.J. and Korach, K.S. (2006) Estrogen receptors and human disease. *J Clin Invest*, **116**, 561-570.
29. Jensen, E.V. and DeSombre, E.R. (1973) Estrogen-receptor interaction. *Science*, **182**, 126-134.
30. Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.-M., Argos, P. and Chambon, P. (1986) Human oestrogen receptor cDNA: Sequence, expression, and homology to v-erbA. *Nature*, **83**, 835-839.
31. Kuiper, G., Enmark, E., Peltö-Huikko, M., Nilsson, S. and Gustafsson, J.A. (1996) Cloning of a novel receptor expressed in rat prostate and ovary. *Proc. Natl. Acad. Sci. USA*, **93**, 5925-5930.
32. Mosselman, S., Polman, J. and Dijkema, R. (1996) ER beta: identification and characterization of a novel human estrogen receptor. *FEBS Lett*, **392**, 49-53.

33. Mader, S., Leroy, P., Chen, J.Y. and Chambon, P. (1993) Multiple parameters control the selectivity of nuclear receptors for their response elements. Selectivity and promiscuity in response element recognition by retinoic acid receptors and retinoid X receptors. *J. Biol. Chem.*, **268**, 591-600.
34. Mader S, Chambon P and JH, W. (1993) Defining a minimal estrogen receptor DNA binding domain. *Nucleic Acids Res*, **21**, 1125-1132.
35. Mader S, Kumar V, de Verneuil H and P, C. (1989) Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. *Nature*, **338**, 271-274.
36. Sanchez, R., Nguyen, D., Rocha, W., White, J.H. and Mader, S. (2002) Diversity in the mechanisms of gene regulation by estrogen receptors. *Bioessays*, **24**, 244-254.
37. Brzozowski, A.M., Pike, A.C., Dauter, Z., Hubbard, R.E., Bonn, T., Engstrom, O., Ohman, L., Greene, G.L., Gustafsson, J.A. and Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, **389**, 753-758.
38. McDonnell, D.P., Clemm, D.L., Hermann, T., Goldman, M.E. and Pike, J.W. (1995) Analysis of estrogen receptor function in vitro reveals three distinct classes of antiestrogens. *Mol Endocrinol*, **9**, 659-669.
39. McInerney, E.M., Weis, K.E., Sun, J., Mosselman, S. and Katzenellenbogen, B.S. (1998) Transcription activation by the human estrogen receptor subtype beta (ER beta) studied with ER beta and ER alpha receptor chimeras. *Endocrinology*, **139**, 4513-4522.
40. McInerney, E.M. and Katzenellenbogen, B.S. (1996) Different regions in activation function-1 of the human estrogen receptor required for antiestrogen- and estradiol-dependent transcription activation. *J Biol Chem*, **271**, 24172-24178.
41. Burris, T. (2001) *The nuclear receptor superfamily*. academic press.
42. Lannigan DA. (2003) Estrogen receptor phosphorylation. *Steroids*, **68**, 1-9.
43. Dutertre M and CL, S. (2003) Ligand-independent interactions of p160/steroid receptor coactivators and CREB-binding protein (CBP) with estrogen receptor-

- alpha: regulation by phosphorylation sites in the A/B region depends on other receptor domains. *Mol Endocrinol*, **17**, 1296-1314. Epub 2003 Apr 1224.
44. Cui, Y., Zhang, M., Pestell, R., Curran, E.M., Welshons, W.V. and Fuqua, S.A. (2004) Phosphorylation of estrogen receptor alpha blocks its acetylation and regulates estrogen sensitivity. *Cancer Res*, **64**, 9199-9208.
 45. Reid, G., Denger, S., Kos, M. and Gannon, F. (2002) Human estrogen receptor-alpha: regulation by synthesis, modification and degradation. *Cell Mol Life Sci*, **59**, 821-831.
 46. Kim, M.Y., Woo, E.M., Chong, Y.T., Homenko, D.R. and Kraus, W.L. (2006) Acetylation of estrogen receptor alpha by p300 at lysines 266 and 268 enhances the deoxyribonucleic acid binding and transactivation activities of the receptor. *Mol Endocrinol*, **20**, 1479-1493.
 47. Wang, C., Fu, M., Angeletti, R.H., Siconolfi-Baez, L., Reutens, A.T., Albanese, C., Lisanti, M.P., Katzenellenbogen, B.S., Kato, S., Hopp, T. *et al.* (2001) Direct Acetylation of the Estrogen Receptor alpha Hinge Region by p300 Regulates Transactivation and Hormone Sensitivity. *J. Biol. Chem.*, **276**, 18375-18383.
 48. Lupien, M., Jeyakumar, M., Hebert, E., Hilmi, K., Cotnoir-White, D., Loch, C., Auger, A., Dayan, G., Pinard, G.A., Wurtz, J.M. *et al.* (2007) Raloxifene and ICI182,780 Increase Estrogen Receptor- α Association with a Nuclear Compartment via Overlapping Sets of Hydrophobic Amino Acids in Activation Function 2 Helix 12. *Mol Endocrinol*, **21**, 797-816.
 49. Sentis, S., Le Romancer, M., Bianchin, C., Rostan, M.C. and Corbo, L. (2005) Sumoylation of the estrogen receptor alpha hinge region regulates its transcriptional activity. *Mol Endocrinol*, **19**, 2671-2684.
 50. Schwabe, J.W., Chapman, L., Finch, J.T. and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*, **75**, 567-578.

51. Hyder SM, Chiappetta C and GM, S. (1999) Interaction of human estrogen receptors alpha and beta with the same naturally occurring estrogen response elements. *Biochem Pharmacol*, **57**, 597-601.
52. Horard, B. and Vanacker, J.M. (2003) Estrogen receptor-related receptors: orphan receptors desperately seeking a ligand. *J Mol Endocrinol*, **31**, 349-357.
53. Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M. and Gannon, F. (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell*, **115**, 751-763.
54. Klochendler-Yeivin, A., Muchardt, C. and Yaniv, M. (2002) SWI/SNF chromatin remodeling and cancer. *Curr Opin Genet Dev*, **12**, 73-79.
55. Ichinose, H., Garnier, J.M., Chambon, P. and Losson, R. (1997) Ligand-dependent interaction between the estrogen receptor and the human homologues of SWI2/SNF2. *Gene*, **188**, 95-100.
56. Inoue, H., Furukawa, T., Giannakopoulos, S., Zhou, S., King, D.S. and Tanese, N. (2002) Largest subunits of the human SWI/SNF chromatin-remodeling complex promote transcriptional activation by steroid hormone receptors. *J Biol Chem*, **277**, 41674-41685.
57. Belandia, B., Orford, R.L., Hurst, H.C. and Parker, M.G. (2002) Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *Embo J*, **21**, 4094-4103.
58. Onate, S.A., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (1995) Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science*, **270**, 1354-1357.
59. Voegel, J.J., Heine, M.J., Zechel, C., Chambon, P. and Gronemeyer, H. (1996) TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. *Embo J*, **15**, 3667-3675.
60. Hong, H., Kohli, K., Trivedi, A., Johnson, D.L. and Stallcup, M.R. (1996) GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the

hormone binding domains of steroid receptors. *Proc Natl Acad Sci U S A*, **93**, 4948-4952.

61. Li, H., Gomes, P.J. and Chen, J.D. (1997) RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2. *Proc Natl Acad Sci U S A*, **94**, 8479-8484.
62. Chen, H., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y. and Evans, R.M. (1997) Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell*, **90**, 569-580.
63. Anzick, S.L., Kononen, J., Walker, R.L., Azorsa, D.O., Tanner, M.M., Guan, X.Y., Sauter, G., Kallioniemi, O.P., Trent, J.M. and Meltzer, P.S. (1997) AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science*, **277**, 965-968.
64. Spencer, T.E., Jenster, G., Burcin, M.M., Allis, C.D., Zhou, J., Mizzen, C.A., McKenna, N.J., Onate, S.A., Tsai, S.Y., Tsai, M.J. *et al.* (1997) Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature*, **389**, 194-198.
65. Heery, D.M., Kalkhoven, E., Hoare, S. and Parker, M.G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, **387**, 733-736.
66. Torchia, J., Rose, D.W., Inostroza, J., Kamei, Y., Westin, S., Glass, C.K. and Rosenfeld, M.G. (1997) The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature*, **387**, 677-684.
67. Chen, D., Ma, H., Hong, H., Koh, S.S., Huang, S.M., Schurter, B.T., Aswad, D.W. and Stallcup, M.R. (1999) Regulation of transcription by a protein methyltransferase. *Science*, **284**, 2174-2177.
68. Chen, D., Huang, S.M. and Stallcup, M.R. (2000) Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300. *J Biol Chem*, **275**, 40810-40816.
69. Kang, Y.K., Guermah, M., Yuan, C.X. and Roeder, R.G. (2002) The TRAP/Mediator coactivator complex interacts directly with estrogen receptors alpha

- and beta through the TRAP220 subunit and directly enhances estrogen receptor function in vitro. *Proc Natl Acad Sci U S A*, **99**, 2642-2647.
70. Metivier, R., Reid, G. and Gannon, F. (2006) Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Rep*, **7**, 161-167.
 71. Reid, G., Hubner, M.R., Metivier, R., Brand, H., Denger, S., Manu, D., Beaudouin, J., Ellenberg, J. and Gannon, F. (2003) Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling. *Mol Cell*, **11**, 695-707.
 72. Metivier, R., Penot, G., Carmouche, R.P., Hubner, M.R., Reid, G., Denger, S., Manu, D., Brand, H., Kos, M., Benes, V. *et al.* (2004) Transcriptional complexes engaged by apo-estrogen receptor-alpha isoforms have divergent outcomes. *Embo J*, **23**, 3653-3666.
 73. Carroll, J.S. and Brown, M. (2006) Estrogen receptor target gene: an evolving concept. *Mol Endocrinol*, **20**, 1707-1714.
 74. Jakacka, M., Ito, M., Weiss, J., Chien, P.-Y., Gehm, B.D. and Jameson, J.L. (2001) Estrogen Receptor Binding to DNA Is Not Required for Its Activity through the Nonclassical AP1 Pathway. *J. Biol. Chem.*, **276**, 13615-13621.
 75. O'Lone R, Frith MC, Karlsson EK and U, H. (2004) Genomic targets of nuclear estrogen receptors. *Mol Endocrinol*, **18**, 1859-1875. Epub 2004 Mar 1818.
 76. Kim K, Barhoumi R, Burghardt R and S, S. (2005) Analysis of estrogen receptor alpha-Sp1 interactions in breast cancer cells by fluorescence resonance energy transfer. *Mol Endocrinol*, **19**, 843-854. Epub 2005 Jan 2006.
 77. DeNardo DG, Kim HT, Hilsenbeck S, Cuba V, Tsimelzon A and PH, B. (2005) Global gene expression analysis of estrogen receptor transcription factor cross talk in breast cancer: identification of estrogen-induced/activator protein-1-dependent genes. *Mol Endocrinol*, **19**, 362-378. Epub 2004 Oct 2028.
 78. Moriarty, K., Kim, K.H. and Bender, J.R. (2006) Minireview: estrogen receptor-mediated rapid signaling. *Endocrinology*, **147**, 5557-5563.

79. Wehling, M., Schultz, A. and Losel, R. (2006) Nongenomic actions of estrogens: exciting opportunities for pharmacology. *Maturitas*, **54**, 321-326.
80. Evinger, A.J., 3rd and Levin, E.R. (2005) Requirements for estrogen receptor alpha membrane localization and function. *Steroids*, **70**, 361-363.
81. Marino, M., Ascenzi, P. and Acconcia, F. (2006) S-palmitoylation modulates estrogen receptor alpha localization and functions. *Steroids*, **71**, 298-303.
82. Jacob, J., Sebastian, K.S., Devassy, S., Priyadarsini, L., Farook, M.F., Shameem, A., Mathew, D., Sreeja, S. and Thampan, R.V. (2006) Membrane estrogen receptors: genomic actions and post transcriptional regulation. *Mol Cell Endocrinol*, **246**, 34-41.
83. Filardo, E.J. and Thomas, P. (2005) GPR30: a seven-transmembrane-spanning estrogen receptor that triggers EGF release. *Trends Endocrinol Metab*, **16**, 362-367.
84. Pedram, A., Razandi, M. and Levin, E.R. (2006) Nature of functional estrogen receptors at the plasma membrane. *Mol Endocrinol*, **20**, 1996-2009.
85. Butler, J.E. and Kadonaga, J.T. (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev*, **16**, 2583-2592.
86. Smale, S.T. (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes Dev*, **15**, 2503-2508.
87. Hochheimer, A. and Tjian, R. (2003) Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes Dev*, **17**, 1309-1320.
88. van Driel, R., Fransz, P. and Verschure, P. (2003) The eukaryotic genome: a system regulated at different hierarchical levels. *J Cell Sci*, **116**, 4067-4075.
89. Turner, B. (2002) Cellular memory and the histone code. *Cell*, **111**, 285-291.
90. Khorasanizadeh, S. (2004) The nucleosome: from genomic organization to genomic regulation. *Cell*, **116**, 259-272.
91. Mellor, J. (2005) The dynamics of chromatin remodeling at promoters. *Mol Cell*, **19**, 147-157.

92. Mellor, J. (2006) Dynamic nucleosomes and gene transcription. *Trends Genet*, **22**, 320-329.
93. Dillon, N. (2006) Gene regulation and large-scale chromatin organization in the nucleus. *Chromosome Res*, **14**, 117-126.
94. Quina, A.S., Buschbeck, M. and Di Croce, L. (2006) Chromatin structure and epigenetics. *Biochem Pharmacol*, **72**, 1563-1569.
95. Grewal, S. and Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science*, **301**, 798-802.
96. Elgin, S.C. (1996) Heterochromatin and gene regulation in *Drosophila*. *Curr Opin Genet Dev*, **6**, 193-202.
97. Zhimulev, I.F., Belyaeva, E.S., Bgatov, A.V., Baricheva, E.M. and Vlassova, I.E. (1988) Cytogenetic and molecular aspects of position effect variegation in *Drosophila melanogaster*. II. Peculiarities of morphology and genetic activity of the 2B region in the T(1;2)dorvar7 chromosome in males. *Chromosoma*, **96**, 255-261.
98. Hager, G.L., Elbi, C., Johnson, T.A., Voss, T., Nagaich, A.K., Schiltz, R.L., Qiu, Y. and John, S. (2006) Chromatin dynamics and the evolution of alternate promoter states. *Chromosome Res*, **14**, 107-116.
99. Woodcock, C.L. (2006) Chromatin architecture. *Curr Opin Struct Biol*, **16**, 213-220.
100. Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R. and Flavell, R.A. (2005) Interchromosomal associations between alternatively expressed loci. *Nature*, **435**, 637-645.
101. Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W. *et al.* (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*, **36**, 1065-1071.
102. Topalidou, I. and Thireos, G. (2003) Gcn4 occupancy of open reading frame regions results in the recruitment of chromatin-modifying complexes but not the mediator complex. *EMBO Rep*, **4**, 872-876.

103. Martone, R., Euskirchen, G., Bertone, P., Hartman, S., Royce, T.E., Luscombe, N.M., Rinn, J.L., Nelson, F.K., Miller, P., Gerstein, M. *et al.* (2003) Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci U S A*, **100**, 12247-12252.
104. Euskirchen, G., Royce, T.E., Bertone, P., Martone, R., Rinn, J.L., Nelson, F.K., Sayward, F., Luscombe, N.M., Miller, P., Gerstein, M. *et al.* (2004) CREB Binds to Multiple Loci on Human Chromosome 22. *Mol. Cell. Biol.*, **24**, 3804-3814.
105. Shang, Y., Myers, M. and Brown, M. (2002) Formation of the androgen receptor transcription complex. *Mol Cell*, **9**, 601-610.
106. Brown, K.E. (2003) Chromatin folding and gene expression: new tools to reveal the spatial organization of genes. *Chromosome Res*, **11**, 423-433.
107. Mahmoudi, T., Katsani, K.R. and Verrijzer, C.P. (2002) GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *Embo J*, **21**, 1775-1781.
108. Chambeyron, S. and Bickmore, W.A. (2004) Does looping and clustering in the nucleus regulate gene expression? *Curr Opin Cell Biol*, **16**, 256-262.
109. Blanton, J., Gaszner, M. and Schedl, P. (2003) Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev*, **17**, 664-675.
110. Zhao, K., Hart, C.M. and Laemmli, U.K. (1995) Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell*, **81**, 879-889.
111. Golovnin, A., Birukova, I., Romanova, O., Silicheva, M., Parshikov, A., Savitskaya, E., Pirrotta, V. and Georgiev, P. (2003) An endogenous Su(Hw) insulator separates the yellow gene from the Achaete-scute gene complex in *Drosophila*. *Development (Cambridge, England)*, **130**, 3249-3258.
112. Byrd, K. and Corces, V.G. (2003) Visualization of chromatin domains created by the gypsy insulator of *Drosophila*. *The Journal of cell biology*, **162**, 565-574.
113. Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing chromosome conformation. *Science*, **295**, 1306-1311.

114. Kleinjan, D.A. and van Heyningen, V. (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet*, **76**, 8-32.
115. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F. and Fraser, P. (2002) Long-range chromatin regulatory interactions in vivo. *Nat Genet*, **32**, 623-626.
116. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell*, **10**, 1453-1465.
117. Dean, A. (2006) On a chromosome far, far away: LCRs and gene expression. *Trends Genet*, **22**, 38-45.
118. Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M. and Fisher, A.G. (1997) Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell*, **91**, 845-854.
119. Iborra, F.J., Pombo, A., Jackson, D.A. and Cook, P.R. (1996) Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci*, **109 (Pt 6)**, 1427-1436.
120. Li, Y.J., Fu, X.H., Liu, D.P. and Liang, C.C. (2004) Opening the chromatin for transcription. *Int J Biochem Cell Biol*, **36**, 1411-1423.
121. Donegan, W. and Spratt, J. (2002) *Cancer of the Breast, Fifth edition*.
122. Kushner, P.J., Agard, D.A., Greene, G.L., Scanlan, T.S., Shiau, A.K., Uht, R.M. and Webb, P. (2000) Estrogen receptor pathways and AP-1. *J Steroid Biochem Mol Biol*, **74**, 311-317.
123. Safe, S. (2001) Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions. *Vitam. Horm.*, **62**, 231-252.
124. Kelly, M.J. and Levin, E.R. (2001) Rapid actions of plasma membrane estrogen receptors. *Trends Endocrinol Metab*, **12**, 152-156.
125. McEwen, B., Akama, K., Alves, S., Brake, W.G., Bulloch, K., Lee, S., Li, C., Yuen, G. and Milner, T.A. (2001) Tracking the estrogen receptor in neurons: implications for estrogen-induced synapse formation. *Proc Natl Acad Sci U S A*, **98**, 7093-7100.

126. Moggs, J.G. and Orphanides, G. (2001) Estrogen receptors: orchestrators of pleiotropic cellular responses. *EMBO Rep*, **2**, 775-781.
127. Agarwal, S. and Rao, A. (1998) Long-range transcriptional regulation of cytokine gene expression. *Curr Opin Immunol*, **10**, 345-352.
128. Wasserman, W.W. and Sandelin, A. (2004) Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet*, **5**, 276-287.
129. Bourdeau, V., Deschenes, J., Metivier, R., Nagai, Y., Nguyen, D., Bretschneider, N., Gannon, F., White, J.H. and Mader, S. (2004) Genome-wide identification of high-affinity estrogen response elements in human and mouse. *Mol Endocrinol*, **18**, 1411-1427.
130. O'Lone, R., Frith, M.C., Karlsson, E.K. and Hansen, U. (2004) Genomic targets of nuclear estrogen receptors. *Mol Endocrinol*, **18**, 1859-1875.
131. Jin, V.X., Sun, H., Pohar, T.T., Liyanarachchi, S., Palaniswamy, S.K., Huang, T.H. and Davuluri, R.V. (2005) ERTargetDB: an integral information resource of transcription regulation of estrogen receptor target genes. *J Mol Endocrinol*, **35**, 225-230.
132. Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F. *et al.* (2006) Genome-wide analysis of estrogen receptor binding sites. *Nat Genet*.
133. Rae, J.M., Johnson, M.D., Scheys, J.O., Cordero, K.E., Larios, J.M. and Lippman, M.E. (2005) GREB 1 is a critical regulator of hormone dependent breast cancer growth. *Breast cancer research and treatment*, **92**, 141-149.
134. Oestreich, K.J., Cobb, R.M., Pierce, S., Chen, J., Ferrier, P. and Oltz, E.M. (2006) Regulation of TCRbeta gene assembly by a promoter/enhancer holocomplex. *Immunity*, **24**, 381-391.
135. Sipos, L. and Gyurkovics, H. (2005) Long-distance interactions between enhancers and promoters. *Febs J*, **272**, 3253-3259.

136. Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F. and de Laat, W. (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet*, **35**, 190-194.
137. de Laat, W. and Grosveld, F. (2003) Spatial organization of gene expression: the active chromatin hub. *Chromosome Res*, **11**, 447-459.
138. Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R. *et al.* (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell*, **122**, 33-43.
139. Papworth, M., Kolasinska, P. and Minczuk, M. (2006) Designer zinc-finger proteins and their applications. *Gene*, **366**, 27-38.
140. Dobrzycka, K.M., Townson, S.M., Jiang, S. and Oesterreich, S. (2003) Estrogen receptor corepressors -- a role in human breast cancer? *Endocr Relat Cancer*, **10**, 517-536.
141. Ratajczak, T. (2001) Protein coregulators that mediate estrogen receptor function. *Reprod Fertil Dev*, **13**, 221-229.
142. Dauvois, S., White, R. and Parker, M.G. (1993) The antiestrogen ICI 182780 disrupts estrogen receptor nucleocytoplasmic shuttling. *J Cell Sci*, **106 (Pt 4)**, 1377-1388.
143. Parker, M.G. (1993) Action of "pure" antiestrogens in inhibiting estrogen receptor action. *Breast cancer research and treatment*, **26**, 131-137.
144. Peekhaus, N.T., Chang, T., Hayes, E.C., Wilkinson, H.A., Mitra, S.W., Schaeffer, J.M. and Rohrer, S.P. (2004) Distinct effects of the antiestrogen Faslodex on the stability of estrogen receptors-alpha and -beta in the breast cancer cell line MCF-7. *J Mol Endocrinol*, **32**, 987-995.
145. Van Den Bemd, G.J., Kuiper, G.G., Pols, H.A. and Van Leeuwen, J.P. (1999) Distinct effects on the conformation of estrogen receptor alpha and beta by both the antiestrogens ICI 164,384 and ICI 182,780 leading to opposite effects on receptor stability. *Biochem Biophys Res Commun*, **261**, 1-5.

146. Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C. *et al.* (2006) Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research*, **16**, 1299-1309.
147. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet*, **38**, 1348-1354.
148. Eeckhoute, J., Keeton, E.K., Lupien, M., Krum, S.A., Carroll, J.S. and Brown, M. (2007) Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res*, **67**, 6477-6483.
149. Barkhem, T., Haldosen, L.A., Gustafsson, J.A. and Nilsson, S. (2002) pS2 Gene expression in HepG2 cells: complex regulation through crosstalk between the estrogen receptor alpha, an estrogen-responsive element, and the activator protein 1 response element. *Molecular pharmacology*, **61**, 1273-1283.
150. Laganier, J., Deblois, G., Lefebvre, C., Bataille, A.R., Robert, F. and Giguere, V. (2005) From the Cover: Location analysis of estrogen receptor alpha target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc Natl Acad Sci USA*, **102**, 11651-11656.
151. Corbi, N., Libri, V., Fanciulli, M., Tinsley, J.M., Davies, K.E. and Passananti, C. (2000) The artificial zinc finger coding gene 'Jazz' binds the utrophin promoter and activates transcription. *Gene therapy*, **7**, 1076-1083.
152. Rodriguez-Garcia, A., Sanchez-Martin, M., Perez-Losada, J., Perez-Mancera, P.A., Sagrera-Aparisi, A., Gutierrez-Cianca, N., Cobaleda, C. and Sanchez-Garcia, I. (2001) Selective destruction of tumor cells through specific inhibition of products resulting from chromosomal translocations. *Current cancer drug targets*, **1**, 109-119.

Annexe 1 : Rational Desing of an Estrogen Receptor Mutant with Altered DNA binding Specificity

Denis Nguyen, Martine Bail, Genevieve Pesant, Virginie Dupont, Étienne Rouault, **Julie Deschênes**, Walter Rocha, Geneviève Melançon, Sergey V. Steinberg and Sylvie Mader

Publié dans: *Nucleic Acid Research*, 2007 Vol 35 No 10 3465-3477

Bien que ma contribution soit secondaire, cet article s'inscrit dans la première année de ma thèse, où j'ai travaillé à comprendre la fixation du récepteur à l'ADN et à maîtriser les différentes techniques servant à définir la fixation du récepteur. Il permet ainsi de bien introduire les concepts de fixation du récepteur à l'ADN

Rational design of an estrogen receptor mutant with altered DNA-binding specificity

Denis Nguyen¹, Martine Bail^{1,2}, Genevieve Pesant¹, Virginie N. Dupont^{1,2}, Étienne Rouault¹, Julie Deschênes^{1,2}, Walter Rocha^{1,2}, Geneviève Melançon¹, Sergey V. Steinberg¹ and Sylvie Mader^{1,2,*}

¹Biochemistry Department and ²Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128 Succursale Centre Ville, Montréal QC H3C 3J7, Canada

Received November 1, 2006; Revised March 24, 2007; Accepted April 2, 2007

ABSTRACT

Although artificial C2-H2 zinc fingers can be designed to recognize specific DNA sequences, it remains unclear to which extent nuclear receptor C4 zinc fingers can be tailored to bind novel DNA elements. Steroid receptors bind as dimers to palindromic response elements differing in the two central base pairs of repeated motifs. Predictions based on one amino acid—one base-pair relationships may not apply to estrogen receptors (ERs), which recognize the two central base pairs of estrogen response elements (EREs) via two charged amino acids, each contacting two bases on opposite DNA strands. Mutagenesis of these residues, E203 and K210 in ER α , indicated that both contribute to ERE binding. Removal of the electric charge and steric constraints associated with K210 was required for full loss of parental DNA-binding specificity and recognition of novel sequences by E203 mutants. Although some of the new binding profiles did not match predictions, the double mutation E203R-K210A generated as predicted a mutant ER that was transcriptionally active on palindromes of PuGGTCA motifs, but not on consensus EREs. This study demonstrates the feasibility of designing C4 zinc finger mutants with novel DNA-binding specificity, but also uncovers limitations of this approach.

INTRODUCTION

Nuclear receptors form a superfamily of ligand-inducible transcription factors that is characterized by two conserved domains, the DNA-binding domain (DBD) composed of two C4 type zinc fingers, and the ligand-binding domain (LBD), which also contains a dimerization

interface (1,2). Nuclear receptors can bind DNA as homo- and/or heterodimers, and recognize response elements arranged as direct repeats, palindromes or inverted palindromes of conserved motifs (3–5). Each motif is bound by the DBD of a single monomer, the two zinc fingers of the DBD combining into a single structural fold with a DNA recognition helix and variable dimerization interfaces (6,7).

Consensus estrogen response elements (EREs, Figure 1A), which are palindromes of PuGGTCA motifs with a three base-pair spacer (8–10), are bound by estrogen receptors (ERs) with highest affinity *in vitro*. Perfect or imperfect EREs present at promoter-proximal locations (11–13) or, as revealed by genome-wide screens, at large distances from the transcriptional start site of estrogen-regulated genes (14,15), are bound by ER α *in vivo* and mediate regulation of estrogen target genes. Other steroid receptors, including the androgen receptor (AR) and glucocorticoid receptor (GR), also bind palindromes with a three base-pair spacer, but the repeated motifs are PuGNACA sequences (16,17). Non-steroid receptors also recognize PuGGTCA or related motifs, but these motifs are arranged as direct repeats or everted repeats with variable spacing. The affinity and selectivity of nuclear receptors for single PuGGTCA motifs is generally low, but can be increased by receptor-specific recognition of additional 5' flanking bases (5,13). Thus, nuclear receptors have achieved selectivity in DNA recognition while interacting with only two main types of motifs. The fact that few variations have been observed in the base-contacting amino acids of the 48 human nuclear receptors (18) suggests that this type of protein–DNA recognition has been conserved throughout evolution, possibly because it affords the most favorable combination of affinity and selectivity. Interestingly, however, nuclear receptor homologs identified in *Caenorhabditis elegans* offer a considerably wider variety of amino acid composition in the DNA recognition helix. Numerous mutations have also been described in the DBD of some nuclear receptors such as

*To whom correspondence should be addressed. Tel: +1 514 343 7166; Fax: +1 514 343 6843; [REDACTED]

© 2007 The Author(s)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.0/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

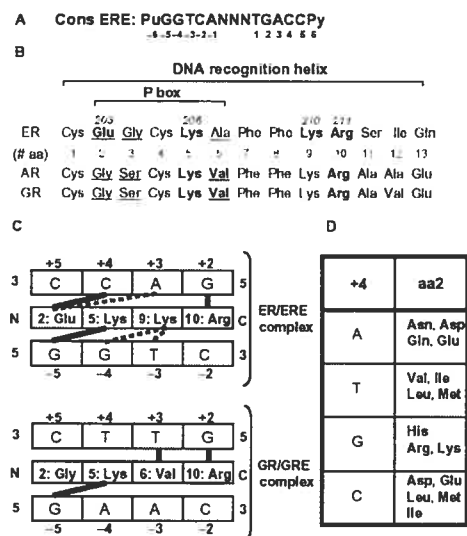


Figure 1. Model for the selectivity of steroid receptor-DNA interactions. (A) Sequence of the consensus ERE and base numbering used in this study. (B) DNA recognition helix of the estrogen (ER), glucocorticoid (GR) and androgen (AR) receptors. The position of amino acids in the DNA recognition helix is indicated by numbers from 1 to 13. The numbering of base-interacting residues in ER α is also indicated. Base-interacting residues in each receptor are in bold, and P box amino acids are underlined. (C) Models of the amino acid-base interactions underlying specific recognition of the estrogen (top) or glucocorticoid (bottom) response element by their cognate steroid receptors. Interactions considered in the model proposed by Suzuki and Yagi (29) are in bold, while interactions described in the ER-ERE crystal structure (7) but not in the model are in dashed lines. (D) Predicted interactions between amino acid replacement at position 2 of the DNA recognition helix and bases at position +4 [from (29)].

VDR and AR, but few changes in DNA-binding patterns have been reported for these mutant receptors (19).

The first clues to the molecular basis of specific DNA recognition by steroid receptors were provided by mutagenesis experiments of the ER and GR, which led to the identification of amino acids responsible for discrimination between the two types of recognition motifs bound by these receptors. Exchanging three amino acids in the ER DBD for the corresponding ones in the GR DBD resulted in a receptor mutant capable of transactivating glucocorticoid target promoters (20). The converse experiment also demonstrated that amino acids at the same three positions (P box, Figure 1B, underlined residues) in the GR were crucial for discriminating between glucocorticoid response elements (GREs) and EREs (21,22). In addition, a loop in the second zinc finger was found to be responsible for specific recognition of two motifs arranged as palindromes with 3 bp spacing (21-23). Crystallographic analyses of complexes between the ER or

GR DBDs and their response elements have uncovered the amino acid base-pair interactions established by these two receptors (7,24). Two residues of the P box, V at the third position in the GR and E at the first position in the ER (E203), contact the central discriminating bases in the ERE and GRE motifs (Figure 1C). In addition, an invariant K residue located further C-terminal in the DNA recognition helix (K210 in ER α) binds the two central bases on the opposite strand with respect to bases contacted by E203 in the ER α ERE complex, but does not participate in contacts in the GR GRE complex (Figure 1C). Other bases common to the ERE and GRE are contacted by amino acids conserved in ER and GR (K206 and R211 in ER α). In addition, these interactions are buttressed by a complex network of amino acid amino acid interactions and amino acid phosphate interactions.

The specificity of response motif recognition by steroid receptors is thus determined by a small number of specific interactions established by 3-4 amino acids. As a consequence, it may be expected that changing the identity of interacting amino acids in the DNA recognition helix would alter the selectivity of receptor DNA interaction as can be achieved with artificial C2H2 zinc fingers (25-28). A simple model for amino acid base interactions within the structural framework of the steroid receptor DBD has been proposed (29). This model relies on chemical rules for possible pairing of amino acid side chains and DNA bases through hydrogen bonding or hydrophobic interactions, and also incorporates stereochemical constraints specific to steroid receptors, based on the position of the DNA recognition helix with respect to the major groove of the DNA. Small, medium or large chains may thus be preferred depending on the position of the interacting amino acids in the DNA recognition helix with respect to the bases (see Figure 1D for possible interactions involving amino acid at position 2). Study of a spontaneous mutation in the first amino acid of the P box of the AR has revealed changes in DNA-binding specificity compatible with the predictions of this model (19). Replacement of G at the first P box position in the AR DBD by R resulted in a mutant that could only bind a subset of the PuGNACA motifs normally bound by this receptor, and this according to the chemical preferences of the R residue with respect to the base at the third position of the motif. However, it remains unclear whether novel types of DNA specificities can be achieved through rational design of nuclear receptor mutants, especially in the case of the ERs, which appear to have more complex determinants of motif recognition than other steroid receptors. The purpose of this study was to examine whether mutating ER α residues that interact with the two central base pairs of EREs either separately or in pairs could generate artificial nuclear receptor DBDs with novel DNA-binding specificities.

MATERIALS AND METHODS

Plasmids

The bacterial expression vector pET3.1-HF81 containing the ER α DBD has been described previously (23).

The bacterial expression vector pET3.1-ER β -DBD was constructed by PCR amplification of a cDNA fragment corresponding to amino acids 140-246 of ER β and subcloning between the KpnI and XhoI sites of a pET3.1. a pET3 (Novagen, San Diego, CA, USA) derivative modified by insertion of a linker containing the KpnI and XhoI sites (23). All ER α DBDs mutants were constructed by insertion between the KpnI and XhoI sites of the pET3.1 vector of a fragment obtained by site-directed mutagenesis of pET3.1-HE81 using PCR amplification.

The wild-type ER α and ER β expression vectors pSG5-HEG0, pCMV-SPORT-ER β have been described previously (15,30). Expression vectors for ER α E203A, ER α E203N, ER α E203H, ER α E203R, ER α K210A, ER α E203A-K210A, ER α E203N-K210A, ER α E203H-K210A, ER α E203R-K210A have been constructed by substitution of the fragment between the KpnI and XhoI sites of pSG1-HE80 (23) by a fragment containing the mutation excised from the corresponding pET3.1 vector.

The tk-CAT reporter plasmids containing one copy of the consensus ERE or of palindromes containing base replacements (pAT-tkCAT, pCT-tkCAT) were derived from the pBL-CAT8+ reporter vector (8) by insertion of double stranded oligonucleotides containing the response elements flanked with BamHI-BglII sites at the BamHI site upstream of the thymidine kinase promoter. The TATA-CAT reporter vectors containing two copies of the consensus ERE or of the CT palindrome were prepared by substitution of the three EREs in pERE3-TATA-CAT (31) by double-stranded oligonucleotides containing two tandem response elements and flanked by BamHI and BglII sites.

Expression in *Escherichia coli* of ER α and ER β DBDs and derivatives thereof

Escherichia coli BL21 DE3 cells were transformed with pET3.1 expression vectors containing the cDNAs for the DBDs of ER α , or of mutants of ER α or ER β and expression was induced in exponentially growing cultures by IPTG (0.5 mM final) for 2 h. Whole bacterial extracts were prepared by sonication in extraction buffer (Tris-HCl pH 7.4, 25 mM; EDTA pH 8.0, 0.1 mM; NaCl 400 mM; glycerol 10%; DTT 1 mM; PMSF 10 mM and protease inhibitors) followed by centrifugation (at 10000g for 15 min) of lysates.

To determine the levels of expression of ER DBDs, aliquots (1 ml) were taken from each culture before induction with IPTG. Bacteria were isolated by centrifugation and resuspended in M9 medium containing each amino acid except methionine and cysteine (0.01% weight/volume each). Rifampicin was added (200 μ g/ml final) to inhibit bacterial RNA polymerase and expression of the T7 polymerase was induced with IPTG (0.5 mM final) for 30 min. [35 S]-methionine was then added and bacteria were further incubated at 37°C for 5 min, collected by centrifugation, resuspended in Laemmli buffer and boiled for 5 min. Labeled proteins were analyzed by electrophoresis on 12% polyacrylamide SDS gel and visualized by fluorography.

Cell culture and transfection

HeLa cells were maintained in DMEM supplemented with 5% fetal bovine serum (FBS). Cells were switched 3 days before transient transfection to medium without phenol red containing 5% FBS pretreated with activated charcoal to remove traces of hormones. For gel shift assays, cells were transiently transfected with ER expression vectors (5 μ g completed to 15 μ g with carrier DNA in 10 cm dish) using the calcium phosphate method (32). For CAT assays, HeLa cells were electroporated (10^7 cells, 0.24 kV, 950 μ F in a Bio-Rad Gene Pulser II apparatus) with varying amounts of expression vectors for wt ER α or for different ER α mutants, 4 μ g of pCMV- β Gal, and 4 μ g of tk-CAT reporter vectors containing single copies of different palindromic response elements. DNA mixes were completed to 80 μ g with salmon sperm DNA in a final volume of 100 μ l of NaCl 210 mM. Electroporated cells were plated in duplicate for parallel immunoblot and CAT assays.

Gel shift assays

Two days post-transfection, HeLa cells were treated for 1 h with 25 nM estradiol and whole cell extracts were prepared by three cycles of freeze-thawing in extraction buffer (Tris HCl pH 7.6, 20 mM; glycerol 20%; KCl 500 mM; DTT 1 mM; EDTA 0.1 mM; PMSF 10 mM and protease inhibitors) followed by centrifugation for 10 min at 13000g. Cell extracts were diluted to a final KCl concentration of 125 mM. Samples were preincubated with 2 μ g poly(dIdC) for 15 min on ice before addition of [32 P]-labeled double-stranded oligonucleotide probes (300 000 c.p.m. per sample). The consensus ERE used for gel retardation assays is derived from the *Xenopus* vitellogenin A2 gene. Reactions were incubated at 25°C for 15 min then loading buffer (0.1% bromophenol blue, 60% glycerol) was added (1/5 V/V). Complexes were separated by electrophoresis on 5% polyacrylamide gels in 0.25 \times TBE (45 mM Tris HCl, 45 mM boric acid and 1 mM EDTA) and visualized by autoradiography. Amounts of bound and free probe were quantified using a Phosphorimaging screen and the Quantity One software from Bio-Rad.

Gels shift assays with whole bacterial extracts containing the ER α or ER β DBD were performed as described above except that extracts were diluted to a final NaCl concentration of 80 mM and 7% polyacrylamide gels in 0.25 \times TBE were used to separate the complexes.

CAT assays

Immediately after seeding, cells were treated with estradiol (25 nM) or vehicle (EtOH) for 40 h. Cells were then harvested and extracts were prepared by three cycles of freeze thawing in CAT extraction buffer (Tris HCl pH 8.0, 0.25 M; DTT 1 mM and protease inhibitors). CAT assays were performed and standardized over β -galactosidase activity as described (31).

Immunoblot analyses

Protein concentrations in whole cell extracts prepared for gel shift or CAT assays were estimated using a Bradford assay. Proteins (50 µg) in Laemmli buffer were separated by electrophoresis on 8% polyacrylamide SDS gels and transferred onto a PVDF membrane. Blots were incubated in blocking solution (PBS 1X, Tween 20 0.05%, BSA 3%) for 1 h and probed with anti-ERα mouse monoclonal antibody B10 (obtained from Prof. P. Chambon) at dilution 1:5000. Membranes were then washed and incubated with a horseradish peroxidase-coupled secondary antibody and visualized with an ECL detection kit.

Modeling

Modeling was performed interactively, using the InsightII/Discover package (Version 2000, Accelrys Inc., San Diego, CA, USA). The X-ray structure of the ER DBD bound to DNA (7) was used as a starting conformation. Each model was submitted to unrestrained energy minimization using the AMBER forcefield (33) until an energy minimum was reached. The presence or absence of particular pair-wise amino acid base interactions in the final structure was treated as a possibility or impossibility to form a particular interaction in a given structural context. Distance measurements between atoms were performed with InsightII tools using a Silicon Graphics O2 computer.

RESULTS

Patterns of response element binding by isolated ERα and ERβ DBDs are similar to each other and to those of the full-length receptors

Estrogen receptors bind with high affinity the consensus palindromic EREs consisting of two PuGGTCA motifs separated by a three base-pair spacer (Figure 1A). Although natural response elements are often imperfect palindromes (12,13), base-pair replacements usually result in a loss of affinity (12,14). To verify that binding patterns are mainly derived from the DBDs of ERs, we have compared the effect of symmetric substitutions at each position of the consensus ERE on binding by full-length ERs transiently expressed in HeLa cells (ERα and ERβ, Figure 2A and B, respectively) or by the ERα or ERβ DBD (ERα DBD and ERβDBD, Figure 2C and D, respectively). The wild-type ERE was bound by either the full-length ERα receptor or the corresponding DBD with the highest affinity (note the higher degree of free probe depletion with wt ERE, Figure 2A and C, lanes 19). Consistent with our previous observations that the DBDs of steroid receptors are sufficient to discriminate between EREs and GREs (23), ER DBDs selectively bound the array of probes in a pattern similar to that of the full-length receptors, although all probes were less efficiently bound with isolated DBDs. This is consistent with a loss of affinity, but not specificity, resulting from the absence of the strong dimerization interface in the LBD (34). In addition, all replacements introduced in both arms of the ERE reduced binding to the same extent for ERα and

ERβ (compare panels A and B, and C and D), which share a high degree of conservation in their DBDs [90% in region C as defined in (35)]. Although the expression levels of transiently expressed full-length ERα and ERβ could not be compared, note that the DBDs were expressed to similar levels as assessed by [³⁵S]Met incorporation (data not shown).

Point mutations in the consensus ERE destabilize interaction with ERs according to chemical compatibility and steric constraints with interacting amino acids

The DNA recognition helix formed by the C-terminal part of the first zinc finger of ERα contains several basic amino acids involved in contacts with bases in the target motifs (Figure 1B, residues in bold; underlined residues are part of the P box). Nucleotides G (−5) and G (+2) interact with residues at position 5 (K206) and 10 (R211) of the DNA recognition helix, respectively (Figure 1C). In addition, K210 (position 9 in the DNA recognition helix) interacts with both G (−4) and T (−3) through direct and water-mediated contacts (Figure 1C). Thus, basic residues are involved in recognition of all positions in the ERE except −/+6 and −/+1, which are not directly contacted in the crystal structure, although both display a preference for purines (Figure 2 A D, lanes 1 3 and 16 18). These direct interactions involving basic amino acids conform to general chemical rules, with G and T, which present only negatively charged groups in the major groove, being preferred over A, which presents both a positively and a negatively charged group, and C, which contains only a positively charged group, leading to unfavorable electrostatic interactions. Accordingly, replacement of G by C in a single motif at position −4 or −5 was sufficient to abolish binding (Figure 3A, lanes 5 and 8). Of note however, the order of the preferred bases is not identical for each contacting basic amino acid. Most noticeably, replacement in a single motif of G at position +2 by T was sufficient to abolish binding (Figure 3A, lanes 13). Molecular modeling suggests that lack of binding to T at position +2 results from steric hindrance due to the methyl group of T−2, which prevents productive interaction between the amines of R211 and O4 of T+2 (Figure 3B).

The pattern of recognition of bound response elements carrying replacements at positions −4 (G>A>T>C) and −3 (T>G>A>C) also differed from that predicted from charge preference due to interaction with basic amino acid K210 (G>T>A>C). Because E203 interacts with bases at positions +3 and +4, i.e. on the other DNA strand, in the crystal structure (7), we compared the roles of both K210 and E203 in modulating patterns of base recognition. While E203 has an opposite and complementary type of chemical selectivity for bases (C>A>T>G) compared to K210, the complex mode of recognition of the two central base pairs by two amino acids may be expected to result in increased steric constraints. At position −4, replacement in a single motif by T drastically reduced binding (Figure 3A, lane 9) and replacement in both motifs of the ERE completely eliminated complex formation (Figure 2A D, lanes 9) and transactivation [data not

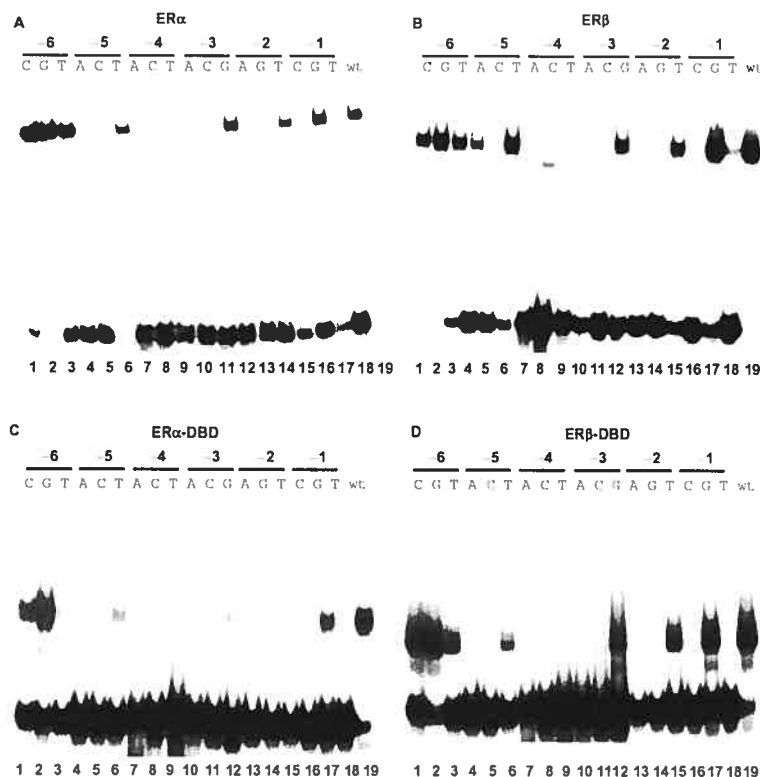


Figure 2. Effect of replacements in both arms of the ERE palindrome on complex formation with full-length ERs and isolated DBDs. (A and B) Gel shift assays performed with whole cell extracts of HeLa cells transiently transfected with expression vectors for ER α (A) or ER β (B), and a panel of probes corresponding to all base possibilities introduced at each position of the response elements in both arms of the palindrome. (C and D) Gel shift assays performed with bacterial extracts containing the DBDs of ER α (C) or ER β (D), using the same panel of probes as in A and B.

shown, see also (36)]. This contrasts with the capacity of other nuclear receptors such as RAR, RXR and VDR to bind to PuGTTCA motifs [(36,37); see also (5) for a review]. Energy minimization indicates that movement of K210 to avoid the methyl group of T-4 prevents interaction with DNA (Figure 3C). In addition, E203 is not capable of interacting with the amino group of A at position +4 (not shown). At position -3, G was less favorable than T (compare lanes 12 and 19 in Figure 3A and Figure 2A–D), contrary to what is observed at the other positions recognized by basic residues (Figure 3A). Modeling indicates that C+3 would create packing problems with E203, preventing interaction of this amino acid with C+4, and K210 is too distant to interact with G-3 (not shown). Thus, our experimental and

modeling data suggest that both E203 and K210 contribute to the selectivity of response element recognition with respect to the two central bases of the ERE.

Roles of E203 and K210 in the specificity and affinity of ER interaction with palindromic response elements

To better analyze the respective roles of E203 and K210 in determining the specificity and affinity of ER interaction with response elements, we further characterized the ER-binding specificity with respect to the two central base pairs of the palindrome motifs using a panel of 16 probes representing all combinations of these central bases introduced in both arms of the palindrome. Full-length ERs and isolated DBDs bound with high affinity only to

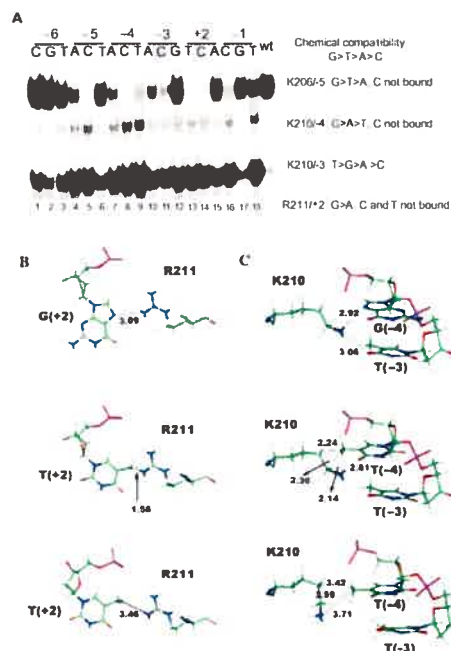


Figure 3. Chemical and steric constraints determine the specificity of complex formation with wt ER α . (A) Gel shift assays performed with the isolated ER α DBD and a panel of probes containing all possible base replacements introduced at each position of a single motif of the consensus ERE. Predicted and observed patterns of interactions are described on the right-hand side of the figure. (B) Interaction between R211 and G+2 as seen in the crystal structure of the ERDBD ERE (7) (top panel). Note the H bond formed between the amino group of R211 and the N7 atom of G+2. Replacement of G+2 (top panel) by T+2 (middle panel) results in steric conflict between the amino groups of R211 and the methyl group of T+2; energy minimization shows that steric hindrance can be relieved, but R211 cannot interact in the major groove with O4 of T+2. (C) Interaction between K210 and G-4 and T-3 as seen in the crystal structure of the ERDBD ERE complex (24) (top panel); note the H bonds formed between the amino group of K210 and the N7 atom of G-4 and the O4 atom of T-3; replacement of G-4 (top panel) by T-4 (middle panel) results in steric conflicts between the side chain of K210 and the methyl group of T-4; energy minimization shows that displacement of the K210 side chain to accommodate this methyl group prevents K210 from interacting with charged groups of T-4 and T-3.

the GT (response elements are designated by bases found on the minus strand at position -4/-3) combination found in the wt ERE and, with lower affinity, to the GG element (Figure 4A–D). Note that complexes formed on element CC with full-length receptors (lane 6) were also observed with extracts from cells transfected with the parental expression vector, and that the smears observed in some lanes with bacterially expressed ER DBDs also appeared with extracts from bacteria transformed with the parental expression vector (not shown).

The very high selectivity of ERs for the two central base pairs in their response elements differs from that of the GR or AR, whose DBDs could both recognize all four PuGNACA-containing palindromes [(20) and data not shown]. Crystallographic analysis of the GR DBD indicates that contacts are established only with the base at position +3 through a V residue (position 6 in the DNA-binding helix). E203 is replaced by a G residue in GR and AR, and the K residue corresponding to K210 in ER α does not contribute to GRE recognition. The higher selectivity of ER interaction with response elements could therefore result either from the fact that E203 interacts with both adjacent +4/+3 bases, and/or from the contribution of K210 in recognizing the -4/-3 bases on the opposite strand of DNA.

To investigate the respective roles of these two amino acids to the affinity and/or selectivity of binding, we replaced either E203 or K210 by A residues in ER α and examined complex formation with the panel of 16 probes corresponding to all possible variants at the two central positions of the repeated motifs. Both mutants still bound with highest efficiency to the consensus ERE (Figure 5A, lanes 12), although the intensity of the retarded band was ~5- to 10-fold lower than with the wt ER α (compare Figure 5A, lanes 12 and 17; see also Figure 5B). However, the K210A mutation was active in transactivation assays using a reporter gene containing the consensus ERE (Figure 8C). Intriguingly, peak transcriptional activity was ~60% higher for K210A than for the wt receptor, suggesting that K210 plays a negative role in transcriptional activation.

Mutant E203A bound with reduced efficacy to the consensus ERE compared to the wt ER α (Figure 5A and B), and formed a weak complex with the PuGATCA probe (AT, Figure 5A, lane 4), which was not detected with the K210A mutation. Finally, the double mutant K210A E203A bound very weakly to the consensus ERE probe, and also formed a detectable complex with the AT probe (Figure 5A, lanes 4, 8, 12 and 16). Binding of E203A to the AT probe suggests that the main role in restricting binding to this element in the wt receptor is played by E203 rather than K210. The absence of high affinity complexes observed with the K210A mutant suggests that the role of K210 in the selectivity of response element recognition is redundant with the role played by E203.

The effects of mutations of E203 on selective DNA binding are not predictable from simple chemical and steric compatibility rules

A simple DNA recognition model for steroid receptors has been proposed previously (8), based on both the general rules of chemical compatibility between amino acids and base pairs, and stereochemical constraints due to the position of the DNA-binding helix across the major groove as derived from the crystal structures of the ER and GR DBDs (Figure 1D for predicted interactions at amino acid position 2). Note that this model is based on a one amino acid-one base interaction relationship, and in particular does not take into account recognition of the

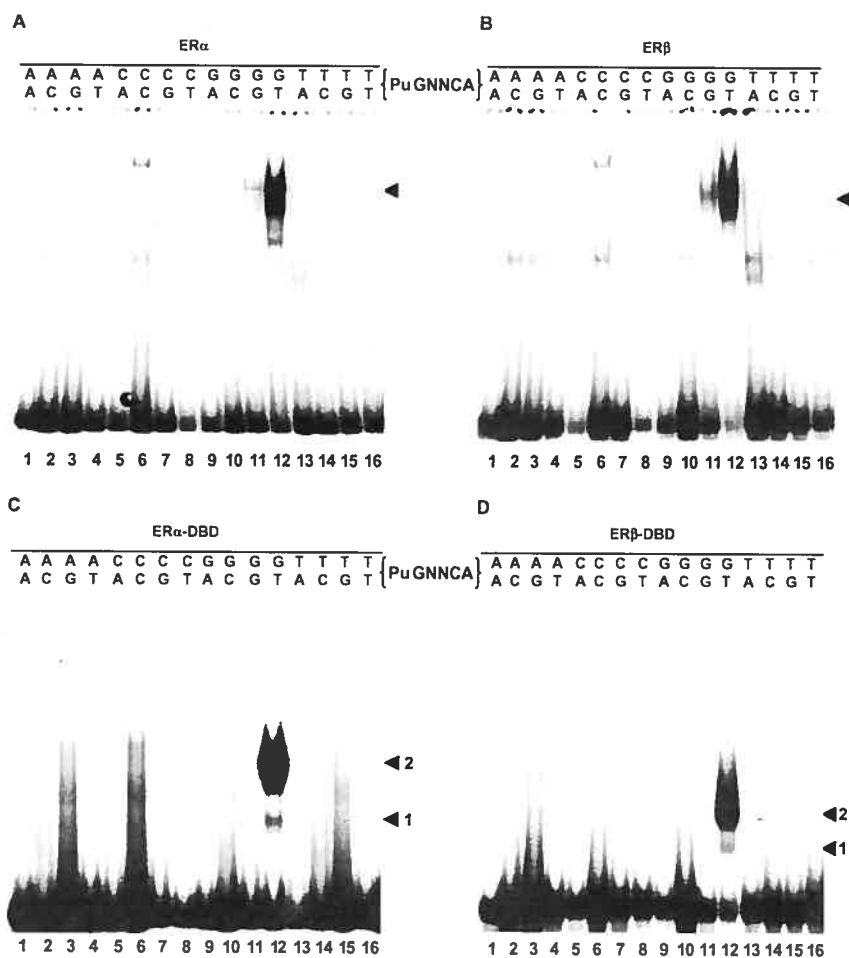


Figure 4. ERs bind with high selectivity to PuGNNCA palindromes. (A) and (B) Extracts from transiently transfected HeLa cells expressing full-length ER α (A) or ER β (B) were incubated with 32 P-labeled oligonucleotide probes containing PuGNNCA palindromes and complexes were separated on 5% polyacrylamide gels. The position of the dimeric complex is indicated by an arrow. (C) and (D) Bacterial extracts containing the ER α (C) or ER β (D) DBDs were incubated with the same probes and complexes were separated by electrophoresis on 7% polyacrylamide gels. The position of specific complexes containing monomers (1) or dimers (2) of ER DBDs are indicated in the figure.

two central bases by both K210 and E203 (Figure 1C, dotted lines), considering only the role of E203 with C + 4 (Figure 1C, bold line). Our finding that E203 and K210 cooperate for selective binding to GT palindromes incited us to test the predictions of this model for replacement of E203 by other amino acids. Replacement of E by N is predicted to switch recognition from C at position +4 to A

(TN elements), and replacement by H or R should lead to specific recognition of a G at this position (CN elements, Figure 1D). These mutations were introduced in ER α full length and in the isolated DBD. Similar expression levels were obtained for all full-length mutants and the wild-type receptor as determined by western analysis (data not shown). Whole cell extracts containing the mutant

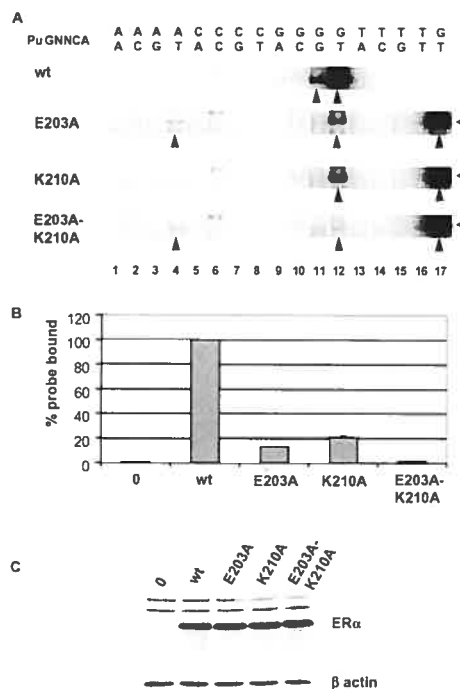


Figure 5. Role of K210 and E203 in the affinity and specificity of receptor DNA complex formation. (A) Gel shift assays performed with wt ER α or the K210A, E203A and E203A-K210A mutants expressed in HeLa cells and a panel of probes containing all possible base combinations at the two central positions of each PuGGTCA motifs in the consensus ERE. Arrows indicate the main specific complexes. Binding of the wt ER α to the consensus ERE (GT) is shown for comparison in lane 17. (B) Quantitation of the levels of binding of the transfected receptors to the consensus ERE probe, expressed as bound over bound plus free probe. Results are an average of three experiments, and error bars indicate SD. (C) Levels of expression of the different transfected receptors as assessed by western analysis using the monoclonal B10 antibody.

receptors E203N and E203H formed complexes mainly with the wt ERE (GT probe, Figure 6A and B, lanes 13 and Figure 6D and E, lanes 12). While binding to the wt ERE was not abolished, the patterns of probe recognition were altered with both mutants, which bound to the AT element (Figure 6A and B, lanes 5 and Figure 6D and E, lanes 4). These interactions are transcriptionally productive, as demonstrated by transient cotransfection of CAT reporter vectors containing the corresponding response elements cloned upstream of the thymidine kinase promoter with expression vectors for the different mutants (Figure 8B, ERE-tk-CAT and AT-tk-CAT

reporter vector). Titration curves were performed using varying concentrations of transfected plasmids, and protein levels were measured by western analysis (Figure 8A). Surprisingly, saturation was reached at identical protein concentrations for all receptors on either response element, indicating that the difference in efficiency of complex formation *in vitro* is not observable in our reporter assay (data not shown). This is possibly due to cooperative effects with other transcription factors bound to the promoter, or to chromatin organization. Peak transactivation levels were similar for the wt and mutant receptors on the consensus ERE, reflecting intact transcription activation properties for the mutants (Figure 8B). On the other hand, mutant E203H, but not E203N, had ~6-fold lower peak levels of transactivation on the AT element than on the consensus ERE (Figure 8B; peak levels on the two response elements were obtained at the same protein concentration). The differential levels of maximal transcriptional activation by the E203H mutant on the two response elements may be related to the previously reported allosteric effect of the DNA sequence on ER α transcriptional activation properties (38,39).

The observed DNA-binding specificity of these mutants do not correspond to predictions based on the proposed model, as no stable binding was observed with TN motifs for E203N (Figure 6A, lanes 14–17), or CN motifs with E203H (Figure 6B, lanes 6–9). Accordingly, the TT and CT elements did not confer estrogen responsiveness to the tk promoter in the presence of these mutants (data not shown). Note that no specific complexes were observed with the E203R mutant on any of the PuGNNCA probes (Figure 6C and F). Lack of binding to the consensus ERE and lack of transactivation on an ERE-tk-CAT reporter (Figure 8C) suggest that R at this position has a destabilizing effect that is stronger than the absence of side chain (A mutation, see Figure 5). Replacement of E203 by R in the structure of the ER-ERE complex reveals that the R side chain is too bulky to fit in the major groove of DNA, and that the amino groups exert repulsive effects with the positively charged groups of C + 4 and C + 5 and steric conflict with C + 4 (Figure 9C and data not shown). In addition, neither binding to CN elements nor transactivation from reporter vectors containing the CT palindrome (Figure 8C, CT-tk-CAT reporter vector) was obtained, indicating again that the predicted switch in specificity has not occurred. These results indicate that a model based on a one amino acid one base-pair relationship is not an accurate description of the interaction between the ER and its target response element, and suggest that K210 plays a role in modulating the DNA-binding specificity of receptors carrying mutations at position 203.

K210 restricts changes in DNA-binding specificity of mutants at position 203

Contrary to the total absence of binding observed with the E203R mutant of ER α , mutation of the corresponding amino acid in the AR (G577) to R led to a different pattern of response element binding than that of the wt receptor. While wt AR bound all four PuGNACA

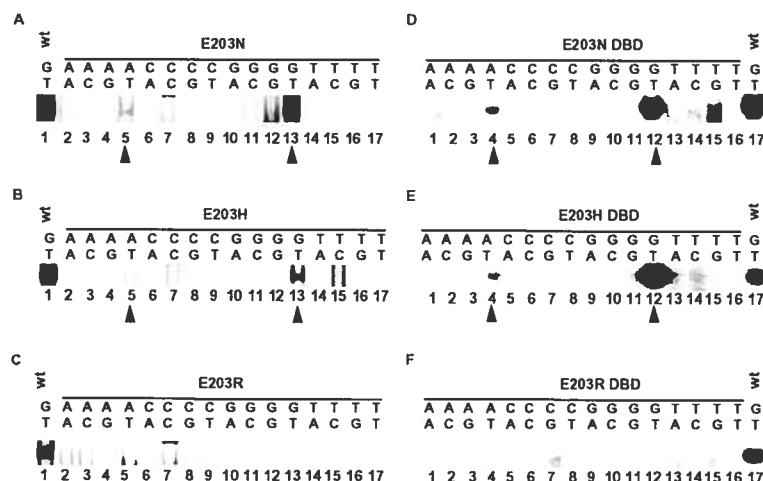


Figure 6. Mutations of E203 do not generate expected switches in DNA-binding specificity. (A, B and C) HeLa cells were transiently transfected with expression vectors for ER α mutant E203N (A), E203H (B) or E203R (C). Whole cell extracts were used in a gel shift assay with the sixteen PuGNNA probes and complexes were separated by electrophoresis on 5% polyacrylamide gels. The complexes formed are indicated by an arrowhead. For comparison, the complex formed between wt ER α and the consensus ERE (GT) is loaded on each gel (lane 1). (D, E and F) Similar assays were performed using bacterial extracts containing mutant ER DBDs carrying the E203N, E203H or E203R mutations.

elements, the G577R mutant bound PuGGACA elements very weakly, but interacted with the three other PuGNACA elements, with a preference for the element containing a G at position +4 (PuGCACA). In the context of the GR DBD, and presumably also the AR DBD, the K residue corresponding to K210 (position 9 in the DNA recognition helix) does not bind DNA, but is involved in an interaction with E at position 13 in the DNA recognition helix (Figure 1B). Thus, we investigated whether mutation K210A may facilitate association of the E203 mutants to novel binding sites and/or reduce binding to the consensus ERE. The resulting E203N/K210A and E203H/K210A double mutants had a much reduced binding to this probe (Figure 7), confirming the role of K210 in binding to the consensus element in the absence of a residue at position 203 recognizing C +4. However, these mutants still did not form stable complexes with probes containing bases predicted to interact with the N or H residue at position 203 (TN, CN, respectively, see Figure 7).

In addition to complexes with AT palindromes already observed with the single mutants, the strongest complexes detected with E203N/K210A and E203H/K210A double mutants were with the AC element (Figure 7). On the other hand, the double mutant E203R-K210A gained binding to the CT element (Figure 7), as could be predicted from chemical compatibility between R at position 203 and G +4 (Figure 1D). This mutant also bound to the AC element (Figure 7). Thus, the specificity of base recognition expected from replacing the

E203 residue by R was revealed in the presence of the K210A mutation, although an additional unpredicted complex was also formed with comparable efficiency. As the level of complex formation on the CT element was much lower than that of the wt receptor on the consensus ERE, we examined whether the interaction between the E203R-K210A mutant and the CT probe is transcriptionally productive. Whereas no estradiol-induced transcription could be detected with wt, K210A or E203R on a CT-tk-CAT reporter vector, the double mutant E203R-K210A gave rise to a detectable increase in estradiol-stimulated transcription (Figure 8C). The double mutant, like the single mutant E203R, was not active on ERE-tk-CAT, confirming the switch in specificity. The complete switch in DNA-binding specificity of the double mutant was confirmed using reporter vectors containing tandem response elements. No transactivation of the reporter vector containing EREs was observed in the presence of the double mutant after estrogen treatment, while peak stimulation of the reporter vector containing CT elements was comparable to that of wt ER α on the reporter containing consensus EREs (Figure 8D).

DISCUSSION

Although C2H2 zinc fingers can be tailored to bind virtually any DNA sequence, nuclear receptors have not demonstrated similar flexibility. A possible reason for success in the rational design of artificial C2H2 zinc finger

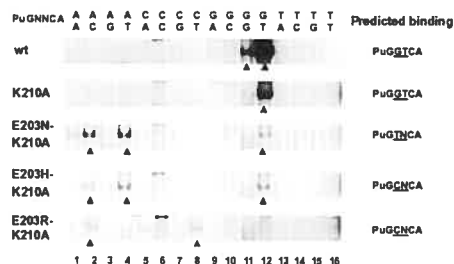


Figure 7. Mutation K210A introduced in E203 mutants reduces binding to the consensus ERE and reveals the switched specificity of mutant E203R. (A) Gel shift assays performed with double mutants of ER α expressed in HeLa cells and the panel of 16 probes containing all possible base combinations at the two central positions of each PuGGTCA motifs in the consensus ERE. Arrows indicate the main specific complexes formed. Note that similar levels of expression were obtained in western analysis (not shown).

proteins is the modular structure of this type of DBDs, each finger recognizing three base pairs and multiple zinc fingers extending the length of bound DNA. On the other hand, steroid receptors bind DNA as dimers recognizing palindromic response elements. As dimerization contributes to the affinity of DNA binding, binding to non-palindromic sequences would likely be of low affinity. Nevertheless, the question remains whether steroid receptors can bind other DNA elements than their two natural cognate response elements. A previously proposed model for prediction of specific amino acid base interactions by the steroid receptor DNA-binding helix (8) incorporates both chemical rules governing amino acid base interactions and stereochemical constraints resulting from the position of the DNA recognition helix in the major groove of response element as determined by crystallographic studies. However, this model relies on one amino acid one base-pair relationships and ignores some of the contacts described in the ER ERE co-crystal structure. In particular, E203 recognizes not only the base at position +4 (Figure 1C, bold line), but also that at position +3 (dotted line), and bases on the opposite strand make contacts with K210. On the other hand, the lysine residue corresponding to K210 in the context of the GR does not bind DNA, and its role in binding of ERs to EREs has not been experimentally confirmed.

Contrary to the GR and AR, which recognize motifs with variable base composition at position -4 conforming to the PuGNACA consensus (20). ERs display a high degree of specificity for bases at position -4/-3. GT was the only element bound with high affinity, while GG was tolerated when introduced in only one motif, such as in the ERE found in the promoter of the rabbit uteroglobin gene (40), but reduced binding efficacy dramatically when introduced in both motifs. This binding pattern is consistent with chemical preferences of E203 for C or A at position +4 and +3, and of K210 for G or T at position -4 and -3, with restrictions imposed by steric constraints

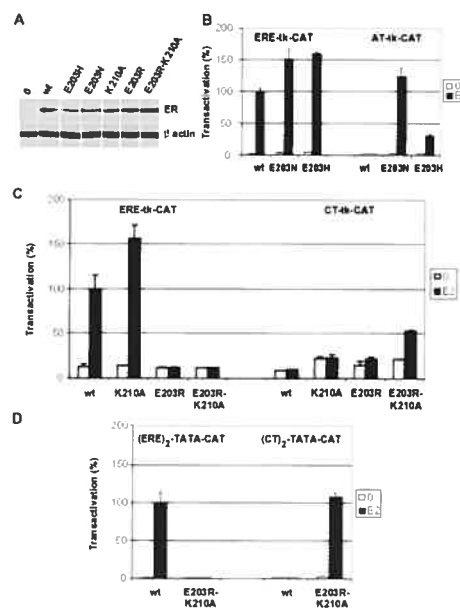


Figure 8. Altered patterns of reporter vector transactivation by ER mutants. HeLa cells were transiently transfected with variable concentrations of expression vectors for the wt ER α or mutants (10 ng–3 μ g per 10 cm plate), tk-CAT reporter vectors containing consensus ERE, AT or CT elements (2 μ g per plate), and an internal control CMV- β Gal reporter vector (2 μ g per plate). Cells were treated with estrogen or vehicle for 40 h and duplicate extracts were assayed for ER protein expression levels in western analysis and for β -galactosidase/CAT activity. (A) Standardized expression levels of wt ER and different mutants. Note that some mutants (E203N and E203R-K210A) were transfected at a 3-fold higher concentration (3 μ g) than other expression vectors (1 μ g) to obtain similar expression levels. (B) CAT activity of wt, E203N and E203H, shown at peak transactivation levels for all plasmids (corresponding to 300 ng for wt and E203H and 1 μ g for E203N). (C) Same as in B, except that expression vectors for wt or the K210A, E203R or E203R-K210A mutants were used together with tk-CAT reporter vectors containing the consensus ERE or the CT element. Peak transactivation was obtained at 300 ng (wt or K210A) or 1 μ g (E203R-K210A). Other plasmids (K210A, E203R), used here at 300 ng, were inactive even at the highest tested concentration (3 μ g, not shown). (D) Same as in C except that expression vectors for the wt ER or the E203R-K210A mutant were used together with TATA-CAT reporter vectors containing two copies of the consensus ERE or the CT element. Maximal transactivation levels, obtained at 1 μ g (wt) or 3 μ g (E203R-K210A), are shown.

with either amino acids. Mutagenesis of each amino acid indicated that both contribute to recognition of the two central base pairs in terms of binding efficiency. E203 also plays a specific role in restricting binding to AT elements, since mutation E203A, but not K210A, allowed formation of complexes. Molecular modeling indicates that this is due to steric conflicts between the carbonyl group of E203 and the methyl group of thymine +4.

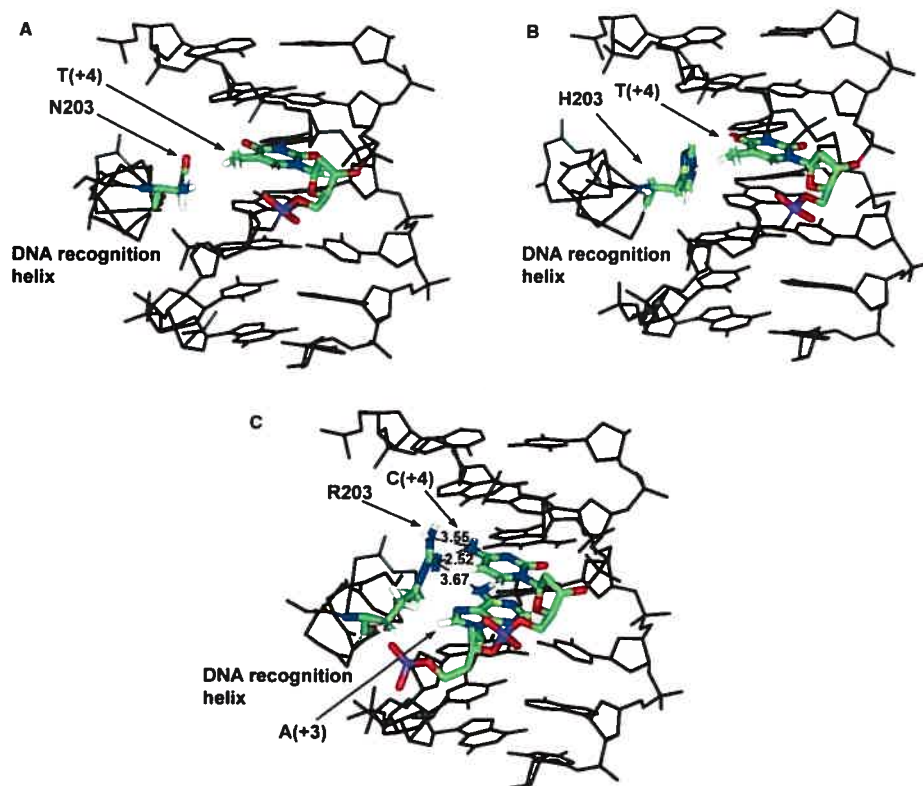


Figure 9. Modeling of the effect of mutations at position 203 on interaction with palindromic response elements. (A and B) Complexes formed between the ER mutant E203N (A) or E203H (B) and the AT element after replacement of E203 by N or H in the ER7DBD ERE crystal structure and energy minimization. The position of the N and H side chains are restrained by the presence of S193. (C) Steric conflicts and charge repulsion with the amino groups of C + 4 in the consensus response element resulting from replacement of E203 by R.

The contribution of K210 to ERE binding likely explains the fact that E203N and E203H mutants still interacted with high affinity with the consensus response elements (GT). Interactions with the AT element, which were not predicted by the proposed model, were observed in gel shift and transactivation assays with both E203N and E203H mutants, and likely result from removal of negative constraints exerted by E203 on T + 4, as observed also in the E203A mutant. Our modeling indicates that the methyl group of thymine can be accommodated by side chain rearrangement when E203 is replaced by N or H (Figure 9A and B). On the other hand, lack of binding of the E203R mutant to all tested elements and of transcriptional activity on the consensus ERE and CT element can be explained by conflicts in charge

preference between R203 and K210 for recognition of the same base pairs.

The role of K210 in preventing a specificity switch by mutations at position 203 is demonstrated by the observation that the double mutant E203R-K210A could bind the CT element, as predicted by base compatibility between R and G + 4, whereas this interaction was not observed with the single mutation E203R. As noted above, the K residue in AR at the corresponding position does not contact DNA, explaining the capacity of the AR mutant with an R at the position equivalent to 203 to interact with G + 4. Another difference between the two receptors is the contribution of V at position 6 in the DNA recognition helix of AR (Figure 1B) in complex stabilization. This may explain the relatively low affinity of the

E203R-K210A mutant for the CT element. However, this interaction was transcriptionally productive. Surprisingly, similar levels of transactivation were obtained with the mutant receptor on the CT element as with the wt receptor on the consensus element at comparable protein concentrations along a range of amounts of transfected plasmids, failing to reveal a different efficiency of reporter DNA saturation (data not shown). Titration curves of mutant E203N and E203H on the consensus ERE and AT element also failed to reveal a shift in protein concentrations necessary to reach peak transcriptional activity. This indicates that our *in vivo* assay may not discriminate between binding sites in the range of affinities that is observable in gel shift assay (within 30-fold of wt ER α -consensus ERE affinity) and may be due to synergism with other transcription factors for binding to our reporter genes *in vivo*. A higher stringency of *in vitro* versus *in vivo* assays is supported by the observation that half-palindromes, which are not bound in our gel shift assay, are enriched, albeit to a much lower degree than high-affinity EREs, in chromatin fragments bound by ER α in a genome-wide analysis (15). On the other hand, our transactivation data confirms the total loss of interaction between the E203R-K210A and the consensus ERE and thus the switch in specificity toward the CT element.

Binding of the E203R-K210A to some unpredicted sites, i.e. the AC element, reflects the limitations of simple models in fully accounting for the complex interactions between these residues and the two central base pairs. In addition, the expected binding of the E203H-K210A mutant to the CN palindromes and of the E203N-K210A mutant to the TN elements were not observed. Replacement of E203 by N or H reveals that both residues are too short to interact with charged groups at position 6 or 7 of A (+4) or G (+4), respectively (Figure 9A and B). Finally, mutants E203H-K210A or E203N-K210A still bound the consensus ERE (GT element), contrary to what was observed with mutation E203R-K210A. The total abrogation of binding to the consensus ERE appears due to charge and steric conflict of R with C + 4 (Figure 9C).

Together, our results indicate that simple chemical and stereochemical rules cannot predict accurately the changes in the selectivity of ER DNA interactions induced by specific mutations in the two central base pairs. A clear limitation is the need to incorporate the contribution of several amino acids to recognition of the same base pairs, and the role of one amino acid in recognizing two adjacent bases. The combined effects of E203 and K210 in interacting with the same bases is apparent both at the level of charge and steric constraints, resulting in the tighter DNA-binding specificity for the two central base pairs observed for ER versus other steroid receptors. Further, steric constraints play an important role in preventing potential interactions. Additional experiments will also be necessary to determine whether the effect of amino acid replacement at other positions in the ER α DNA-binding helix, which are involved in simpler one residue one base interactions, is more easily predictable. It remains possible that other receptors may be more amenable to the rational design of mutant receptors with

altered DNA-binding specificity, due to differences in composition of the DNA-binding helix and/or in the mode of DBD dimerization. In this respect, it will be of interest to investigate the DNA-binding specificity of non-classical *C. elegans* receptors, which contain widely diverging combinations of amino acids in their DNA recognition helix. Finally, combinatorial approaches as performed for C2H2 zinc fingers (41,42) may help clarify how amino acids of the DNA-binding helix cooperate toward the establishment of novel DNA-binding specificities.

ACKNOWLEDGEMENTS

This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) grant to S.M. J.D. has been supported by an FRSQ fellowship, and V.N.D. by a studentship from the Montreal Centre for Experimental Therapeutics in Cancer Canadian Institutes for Health Research teaching program. S.V.S. is a CIHR scientist, and S.M. is an FRSQ Chercheur-Boursier Senior and holds the CIBC Breast Cancer Research Chair at Université de Montréal. Funding to pay the Open Access publication charges for this article was provided by NSERC.

Conflict of interest statement. None declared.

REFERENCES

- Green, S. and Chambon, P. (1988) Nuclear receptors enhance our understanding of transcription regulation. *Trends Genet.*, **4**, 309–314.
- Beato, M. and Klug, J. (2000) Steroid hormone receptors: an update. *Hum. Reprod. Update*, **6**, 225–236.
- Glass, C.K. (1994) Differential recognition of target genes by nuclear receptor monomers, dimers and heterodimers. *Endocr. Rev.*, **15**, 391–407.
- Rastinejad, F., Perlman, T., Evans, R.M. and Sigler, P.B. (1995) Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature*, **375**, 203–211.
- Khorasanzadeh, S. and Rastinejad, F. (2001) Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.*, **26**, 384–390.
- Baumann, H., Paulsen, K., Kovacs, H., Berglund, H., Wright, A.P., Gustafsson, J.A. and Hard, T. (1993) Refined solution structure of the glucocorticoid receptor DNA-binding domain. *Biochemistry*, **32**, 13463–13471.
- Schwabe, J.W., Chapman, L., Finch, J.T. and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*, **75**, 567–578.
- Klein-Hitpass, L., Ryffel, G.U., Heitlinger, F. and Cato, A.C. (1988) A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. *Nucleic Acids Res.*, **16**, 647–663.
- Driscoll, M.D., Sathya, G., Muyan, M., Klinge, C.M., Hilf, R. and Bambara, R.A. (1998) Sequence requirements for estrogen receptor binding to estrogen response elements. *J. Biol. Chem.*, **273**, 29321–29330.
- Kulakosky, P.C., McCarty, M.A., Jernigan, S.C., Risinger, K.E. and Klinge, C.M. (2002) Response element sequence modulates estrogen receptor alpha and beta affinity and activity. *J. Mol. Endocrinol.*, **29**, 137–152.
- Walker, P., Gernoud, J.-E., Brown-Luedi, M., Givel, F. and Wahli, W. (1981) Sequence homologies in the region preceding the transcription initiation site of the liver estrogen-responsive vitellogenin and apo-VLDLII gene. *Nucleic Acids Res.*, **12**, 8611–8626.

12. Klinge, C.M. (2001) Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Res.*, **29**, 2905–2919.
13. Sanchez, R., Nguyen, D., Rocha, W., White, J.H. and Mader, S. (2002) Diversity in the mechanisms of gene regulation by estrogen receptors. *Bioessays*, **24**, 244–254.
14. Bourdeau, V., Deschenes, J., Melivier, R., Nagai, Y., Nguyen, D., Bretschneider, N., Gannon, F., White, J.H. and Mader, S. (2004) Genome-wide identification of high-affinity estrogen response elements in human and mouse. *Mol. Endocrinol.*, **18**, 1411–1427.
15. Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Beckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C. *et al.* (2006) Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.*
16. Klock, G., Strähle, U. and Schütz, G. (1987) Oestrogen and glucocorticoid responsive elements are closely related but distinct. *Nature*, **329**, 734–736.
17. Nordeen, S.K., Suh, B.J., Kühnel, B. and Hutchinson, C.A. (1990) Structural determinants of a glucocorticoid receptor recognition element. *Mol. Endocrinol.*, **4**, 1866–1873.
18. Robinson-Rechavi, M., Escrivá García, H. and Laudet, V. (2003) The nuclear receptor superfamily. *J. Cell Sci.*, **116**, 585–586.
19. Nguyen, D., Steinberg, S.V., Rouault, E., Chagnon, S., Gottlieb, B., Pinsky, L., Trifiro, M. and Mader, S. (2001) A G577R mutation in the human AR P box results in selective decreases in DNA binding and in partial androgen insensitivity syndrome. *Mol. Endocrinol.*, **15**, 1790–1802.
20. Mader, S., Kumar, V., de Verneuil, H. and Chambon, P. (1989) Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. *Nature*, **338**, 271–274.
21. Umesono, K. and Evans, R.M. (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*, **57**, 1139–1146.
22. Danielsen, M., Hinck, L. and Ringold, G.M. (1989) Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. *Cell*, **57**, 1131–1138.
23. Mader, S., Chambon, P. and White, J.H. (1993) Defining a minimal estrogen receptor DNA binding domain. *Nucleic Acids Res.*, **21**, 1125–1132.
24. Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. and Sigler, P.B. (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, **352**, 497–505.
25. Sera, T. and Uranga, C. (2002) Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. *Biochemistry*, **41**, 7074–7081.
26. Blancafort, P., Segal, D.J. and Barbas, C.F.III. (2004) Designing transcription factor architectures for drug discovery. *Mol. Pharmacol.*, **66**, 1361–1371.
27. Mandell, J.G. and Barbas, C.F.III. (2006) Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.*, **34**, W516–523.
28. Papworth, M., Kolasinska, P. and Minczuk, M. (2006) Designer zinc-finger proteins and their applications. *Gene*, **366**, 27–38.
29. Suzuki, M. and Yagi, N. (1994) DNA recognition code of transcription factors in the helix-turn-helix, probe helix, hormone receptor, and zinc finger families. *Proc. Natl. Acad. Sci. USA*, **91**, 12356–12361.
30. Tora, L., Mullick, A., Metzger, D., Ponglikitmongkol, M., Park, I. and Chambon, P. (1989) The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties. *EMBO J.*, **8**, 1981–1986.
31. Barsalou, A., Gao, W., Anghel, S., Carriere, J. and Mader, S. (1998) Estrogen response elements can mediate agonist activity of antiestrogens in human endometrial Ishikawa cells. *J. Biol. Chem.*, **273**, 17138–17146.
32. Banerji, J., Rusconi, S. and Schaffner, W. (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, **27**, 299–308.
33. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, T.E.I., Ferguson, D.M., Seibel, G.L., Singh, U.C., Weiner, P.K. *et al.* (1995) AMBER 4.1 University of California, San Francisco, CA, USA.
34. Brzozowski, A.M., Pike, A.C., Dauter, Z., Hubbard, R.E., Bonn, T., Engström, O., Olsson, L., Greer, G.L., Gustafsson, J.A. *et al.* (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, **389**, 753–758.
35. Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J.M. and Chambon, P. (1986) The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. *EMBO J.*, **5**, 891–897.
36. Mader, S., Leroy, P., Chen, J.Y. and Chambon, P. (1993) Multiple parameters control the selectivity of nuclear receptors for their response elements. Selectivity and promiscuity in response element recognition by retinoic acid receptors and retinoid X receptors. *J. Biol. Chem.*, **268**, 591–600.
37. Mader, S., Chen, J.Y., Chen, Z., White, J., Chambon, P. and Gronemeyer, H. (1993) The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificities of the DNA binding domains. *EMBO J.*, **12**, 5029–5041.
38. Wood, J.R., Greene, G.R. and Nardulli, A.M. (1988) Estrogen response elements function as allosteric modulators of estrogen receptor conformation. *Mol. Cell. Biol.*, **18**, 1927–1934.
39. Hall, J.M., McDonnell, D.P. and Korach, K.S. (2002) Allosteric regulation of estrogen receptor structure, function, and coactivator recruitment by different estrogen response elements. *Mol. Endocrinol.*, **16**, 469–486.
40. Slater, E.P., Redeuilh, G., Theis, K., Suske, G. and Beato, M. (1990) The uteroglobin promoter contains a noncanonical estrogen responsive element. *Mol. Endocrinol.*, **4**, 604–610.
41. Choo, Y. and Klug, A. (1994) Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci. USA*, **91**, 11163–11167.
42. Magnemat, L., Blancafort, P. and Barbas, C.F.III. (2004) *In vivo* selection of combinatorial libraries and designed affinity maturation of polydactyl zinc finger transcription factors for ICAM-1 provides new insights into gene regulation. *J. Mol. Biol.*, **341**, 635–649.